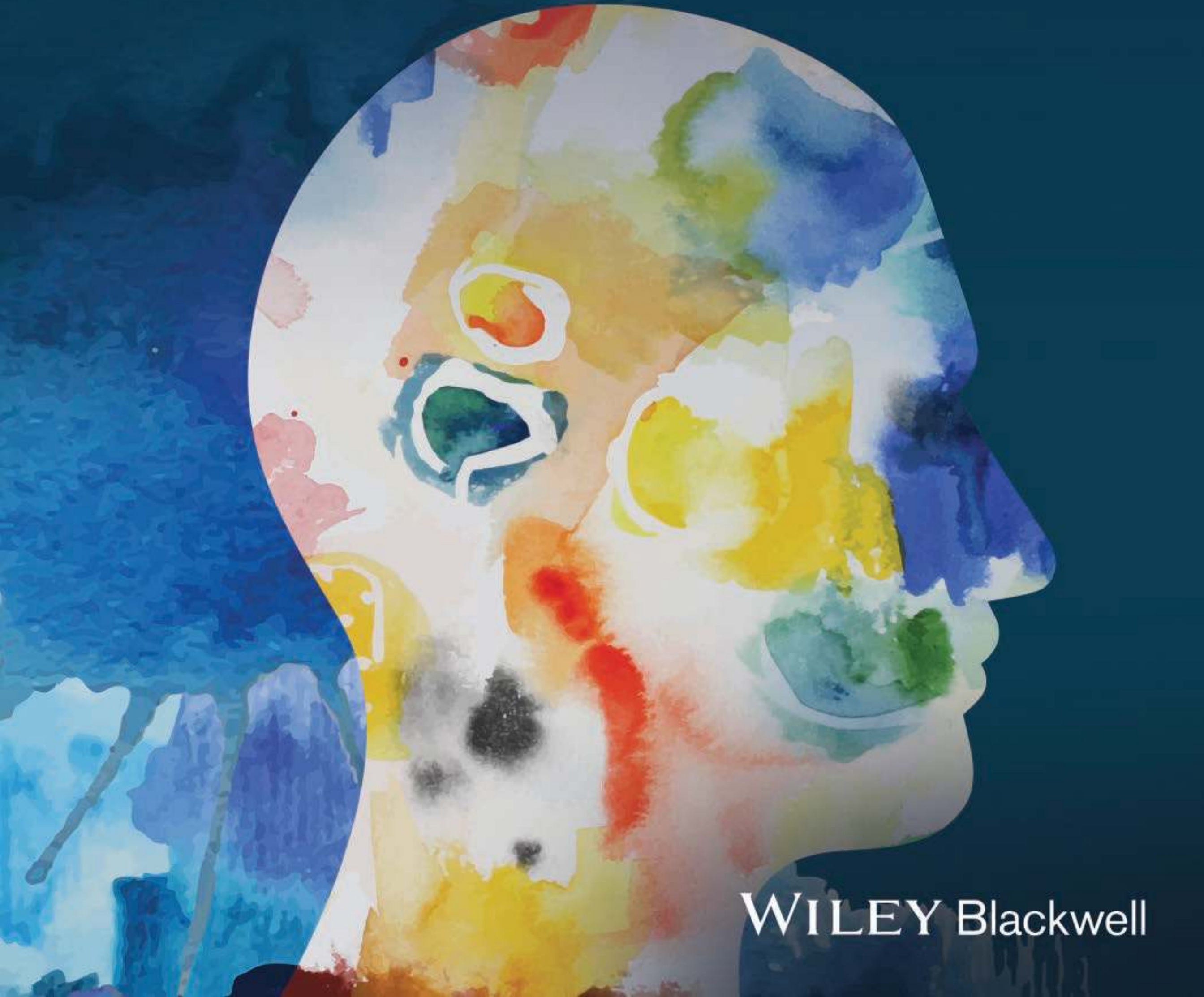
Brandon Rickabaugh
J.P. Moreland

The Substance of Consciousness

A Comprehensive Defense of Contemporary Substance Dualism



WILEY Blackwell

The Substance of Consciousness

The Substance of Consciousness

A Comprehensive Defense of Contemporary Substance Dualism

Brandon Rickabaugh
J. P. Moreland

WILEY Blackwell

Copyright © 2024 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permission.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

A catalogue record for this book is available from the Library of Congress

Hardback ISBN: 9781394195473; Paperback ISBN: 9781394195480; ePub ISBN: 9781394195497;
ePDF ISBN: 9781394195503

Cover image: © DrAfter123/Getty Images
Cover design: Wiley

Set in 10.5/13pt STIXTwoText by Integra Software Services Pvt. Ltd., Pondicherry, India

“There is exactly one overriding question in contemporary philosophy....How do we fit in?...How can we square this self-conception of ourselves as mindful, meaning-creating, free, rational, etc., agents with a universe that consists entirely of mindless, meaningless, unfree, nonrational, brute physical particles?”

John Searle, *Freedom and Neurobiology*

To Richard Swinburne, whose meticulous reformulation and rigorous defense of substance dualism has inspired and challenged generations of philosophers to contend with the reality that we are body and soul.

To the prolific substance dualists of our lifetime who have greatly inspired and informed our work: Todd Buras, Roderick Chisholm (in his better moments), C. Stephen Evans, Stewart Goetz, W. D. Hart, William Hasker, Robert Koons, H. D. Lewis, E. J. Lowe, David Lund, Alexander Pruss, Uwe Meixner, Charles Taliaferro, Dean Zimmerman, and Dallas Willard.

Brief Table of Contents

List of Figures xxi

Acknowledgments xxii

PART I Ontologically Serious Philosophy of Mind 1

1 Substance Dualism in the 21st Century 3

2 How Staunch Naturalism Shapes the Dialectic in Philosophy of Mind 16

3 A Brief Ontological Detour: Subjects, Parts, Wholes, and Unity 43

PART II Arguments from Introspection, Self-Awareness, and Intentionality 57

4 The Real Nature of Introspection Arguments for Substance Dualism 59

5 From Self-Awareness and Intentionality to the Self as Soul 74

PART III On the Fundamental Unity of Conscious Beings 117

6 From Phenomenal Unity to the Synchronic Unity of the Immaterial Self 119

7 Mereological Essentialism and the Diachronic Endurance of the Soul 144

PART IV Updated and Novel Arguments from Modality and Libertarian Freedom 189

8 Upgrading Modal Arguments for Substance Dualism 191

9 Staunch Libertarian Agency and the Simple, Enduring Soul 234

PART V New and Neglected Responses to Common Defeaters Against Substance Dualism 273

10 Important Frequently Raised Defeaters Against Substance Dualism 275

PART VI Charting a Future for Substance Dualism 309

11 New Research Programs for 21st Century Substance Dualism 311

APPENDIX: The Soul, Mental Action, and the Conservation Laws 344
Mihretu P. Guta

Bibliography 361

Index 392

Contents

List of Figures xxi
Acknowledgments xxii

PART I Ontologically Serious Philosophy of Mind 1

CHAPTER 1 Substance Dualism in the 21st Century 3
- 1.1 The Return of Substance Dualism 3
  - 1.1.1 The Great Revolt Against Dualism 3
  - 1.1.2 The Resurgence of Substance Dualism 5
- 1.2 The Case Against Substance Dualism Weakens 6
  - 1.2.1 Admissions to the Weakness of Standard Objections to Substance Dualism 6
  - 1.2.2 Admitting the Faith of Anti-Dualism 6
  - 1.2.3 Proposals of Naturalistic Substance Dualism 7
- 1.3 The Revival of Substance Dualist Intuitions 8
  - 1.3.1 The Return to the Self and Subject of Consciousness 8
  - 1.3.2 The Meta-Problem of Consciousness and Cognitive Science 8
  - 1.3.3 The Revival of Aristotelianism 8
  - 1.3.4 Broad Worldview Considerations 9
- 1.4 The Current Turning Away From Standard Physicalism 9
  - 1.4.1 The Fundamentality of Consciousness 10
  - 1.4.2 The Return of Panpsychism 10
  - 1.4.3 Consciousness-First Philosophy of Mind 11
  - 1.4.4 Phenomenal Transparency and the Opacity of the Physical 11
- 1.5 Mere Substance Dualism 13
  - 1.5.1 Delineating the View 13
  - 1.5.2 The Undeniable Recalcitrance of Substance Dualism 15

CHAPTER 2 How Staunch Naturalism Shapes the Dialectic in Philosophy of Mind 16
- 2.1 Worldviews, Scientific Naturalism, and the Standard Picture 16
  - 2.1.1 Philosophy of Mind and Worldview Implications 16
  - 2.1.2 Two Versions of a Naturalistic Worldview 17
  - 2.1.3 Naturalism and the Standard Picture 18

2.2 Important Epistemic Background Issues 20
  2.2.1 Basicality 20
  2.2.2 Naturalness 20
  2.2.3 Explanatory Simplicity: Ockham's Razor and Ontological Parsimony 21
2.3 Contemporary Scientific Naturalists Should Be Staunch Naturalists and Strong Physicalists 23
  2.3.1 The Naturalist Epistemic Attitude 24
    2.3.1.1 Epistemology and the Origin and Nature of Human Noetic Faculties 24
    2.3.1.2 Scientism 25
    2.3.1.3 Third-Person Ways of Knowing 25
    2.3.1.4 Combinatorial Modes of Explanation 25
  2.3.2 The Grand Story 27
    2.3.2.1 Précis of the Grand Story 27
    2.3.2.2 Four Core Elements to the Grand Story 27
  2.3.3 The Standard Naturalist Ontology 28
    2.3.3.1 The Location Problem 28
    2.3.3.2 Microphysical Priority and the Mereological Hierarchy 29
2.4 Naturalistic Emergence 29
  2.4.1 Three Important Distinctions 30
    2.4.1.1 Weak Vs. Strong Property Supervenience 30
    2.4.1.2 Emergent Vs. Structural Supervenience 30
    2.4.1.3 Making Precise the Notion of an Emergent Property 31
  2.4.2 Additional Metaphysical Aspects of Emergent Properties 31
    2.4.2.1 Twin Hallmarks of Emergence 31
    2.4.2.2 Seven Characteristics of Emergent Properties 32
  2.4.3 Confusion About the Basic Problem of Emergent Properties and Faint-Hearted Naturalism 32
  2.4.4 Classification of Emergent Properties and a New Difficulty for Faint-Hearted Naturalism 34
  2.4.5 Eight Additional Criticisms 35
    2.4.5.1 *"Emergence" Is Just a Label* 35
    2.4.5.2 Emergence and Empirically Equivalent Theories 36
    2.4.5.3 Emergence *Ex Nihilo* 36
    2.4.5.4 Emergence as Brute Facts 37
    2.4.5.5 Sorites Problems 38
    2.4.5.6 The Combination Problem 38
    2.4.5.7 Consciousness and a Bayesian Argument for Theism 39
    2.4.5.8 Novel Properties and New Substances 40
2.5 Going Forward 41

CHAPTER 3 A Brief Ontological Detour: Subjects, Parts, Wholes, and Unity 43
3.1 Husserlian Parts, Wholes, and Unity 43

3.1.1 Parthood and Simplicity: Separable and Inseparable Parts 44
3.1.2 Holistic Unity: Genuine Wholes, Relations, and Aggregates 46
3.2 Subjects of Consciousness: A Few Metaphysical Theses 50
3.2.1 What Is Consciousness? 50
3.2.2 Realism About Substantial Subjects and the Self 53
3.2.3 Subject Necessity 54
3.2.4 Subjects, Not Mental States, Are Conscious 55
3.2.5 Subject Simplicity and Subject Complexity 56

PART II Arguments from Introspection, Self-Awareness, and Intentionality 57

CHAPTER 4 The Real Nature of Introspection Arguments for Substance Dualism 59
4.1 Introspection Arguments for Substance Dualism 59
4.2 The Intensional Fallacy Objection 60
4.2.1 Replies 62
4.2.1.1 Equivocating Propositional Knowledge and Knowledge by Acquaintance 62
4.2.1.2 Absurd Introspection Skepticism 63
4.2.1.3 Missing the Target: Why AFI Is Metaphysical 66
4.3 The Objection from Neurological *de Dicto* but Not *de Re* Beliefs 67
4.3.1 Replies 68
4.3.1.1 The Neuroscientific Loss of Introspective Skepticism 68
4.3.1.2 The Illicit Presumption of Physicalism 70
4.3.1.3 Dubious Analogies from the History of Science 70
4.3.1.4 Grahek's Challenge to the Phenomenal Essence of Pain 71
4.4 Searle's Anti-Introspection Argument 72
4.5 Conclusion 73

CHAPTER 5 From Self-Awareness and Intentionality to the Self as Soul 74
5.1 David Barnett's Simplicity Argument 75
5.1.1 Barnett's Argument Stated 75
5.1.2 Objections to Barnett's Argument 76
5.1.2.1 Elimination 76
5.1.2.2 Maximality 77
5.1.2.3 Substance 78
5.1.2.4 Integrity 79
5.1.3 Summary 82
5.2 The Hard Meta-Problem of Consciousness 82
5.2.1 Intuitions and Rational Seemings 82
5.2.2 The Easy and Hard Meta-Problems of Consciousness 83
5.2.3 The Data: Dualist Seemings 84
5.2.3.1 Explanatory Gap Seemings 84

- 5.2.3.2 Unity and Simplicity Seemings 85
- 5.2.3.3 The Cognitive Science of Commonsense Substance Dualism 86
- 5.2.3.4 Eighteen Distinct Dualist Seemings 88
- 5.2.3.5 Artificial Intelligence Test Seemings 90
- 5.2.3.6 Three Explanatory Conditions of Dualist Seemings 90

5.3 A Direct Self-Awareness Argument for Substance Dualism 92
- 5.3.1 DSA: The Direct Self-Awareness Account of Dualist Seemings 92
- 5.3.2 Knowledge by Acquaintance 94
- 5.3.3 Monadic Intentionality 96
- 5.3.4 Givenness and Directedness 97
- 5.3.5 Direct Self-Awareness as a Self-Presenting Property 98
- 5.3.6 Self-Awareness Necessity 99
- 5.3.7 The Causal-Acquaintance Principle 100
- 5.3.8 Monadic Intentionality, Direct Self-Awareness, and Dualist Seemings 101
- 5.3.9 Satisfying the Explanatory Conditions of Dualist Seemings 101
- 5.3.10 Explanatory Unity 102

5.4 A Direct Reflective Self-Awareness Argument for Substance Dualism 103
- 5.4.1 Husserlian Fulfillment Structures and Self-Knowledge 103
- 5.4.2 Knowledge and Intentionality 103
- 5.4.3 Fulfilled SD Seemings 104

5.5 Objections to Self-Awareness Arguments 105
- 5.5.1 Hume’s Phenomenological Report 105
- 5.5.2 Commonsense Materialism 106
- 5.5.3 Linguistic and Theoretical Dependence 107
- 5.5.4 Absent Dualist Seemings 107
- 5.5.5 Religious Etiology 108
- 5.5.6 Neuroscientific Defeaters 108
- 5.5.7 Self-Awareness Failure 108
- 5.5.8 The Dual-Process Defeater 110
  - 5.5.8.1 Dualist Seemings and Moorean Facts 112
  - 5.5.8.2 A Dualist Rival to DPA 113
  - 5.5.8.3 A Generality Problem for DPA 114
  - 5.5.8.4 A Superior Physicalist Account of Self-Awareness 115

5.6 A Crucial Implication of the Direct Self-Awareness Argument 116

PART III On the Fundamental Unity of Conscious Beings 117

CHAPTER 6 From Phenomenal Unity to the Synchronic Unity of the Immaterial Self 119
- 6.1 Introduction 119
- 6.2 The Datum: Phenomenally Unified Consciousness 121

6.2.1 Phenomenal Unity 121
6.2.2 Phenomenal Holism 122
6.2.3 Subject Phenomenal Unity 122
6.3 Phenomenally Unified Consciousness Contra Physicalism 123
6.3.1 The Anti-Distribution Argument 123
6.3.2 A Temporal Distribution Problem 125
6.3.3 The Process Argument 126
6.3.4 A Non-Phenomenal Processing Problem 127
6.3.5 A Micro-Subject Problem 127
6.3.6 The Function Argument 128
6.3.6.1 The Function Argument Against Animalism 129
6.3.7 An Emergentist Objection 131
6.3.8 Not All Simple Souls Will Do: Contra Emergent Substance Dualism 132
6.4 Phenomenally Unified Consciousness Contra Russellian Panpsychism 133
6.4.1 The Combination Problem(s) 134
6.4.2 The Privacy Argument Against Panpsychist Phenomenal Unity 135
6.4.2.1 Mental-Sharing and Subjectivity 135
6.4.2.2 Self-Presentation 136
6.4.2.3 Mine-ness Necessity 137
6.4.2.4 A Neo-Aristotelian Ontology of Phenomenal States 137
6.4.3 Objections to the Privacy Argument 139
6.4.3.1 Deny Mental-Sharing 139
6.4.3.2 Reject Privacy 139
6.5 General Objections to the Unity of Consciousness Argument 141
6.5.1 Rejecting Phenomenal Unity 141
6.5.2 Tim Bayne's Explanatory Failure Objection 142

CHAPTER 7 Mereological Essentialism and the Diachronic Endurance of the Soul 144
7.1 Stating and Clarifying the Argument 145
7.1.1 The Mereological Argument 145
7.1.2 Simplicity 146
7.1.3 Clarification of the Argument 147
7.2 The Standard Mereological Hierarchy or Layered View of the Natural World 148
7.2.1 A General Depiction of the Mereological Hierarchy 149
7.2.2 The Category of Individual 149
7.2.3 *The Category of Property/Relation* 149
7.2.4 The Structure of the Mereological Hierarchy 150
7.2.5 Causation 153
7.2.6 Concepts and Laws 153
7.3 A Defense of Premise (4): Human Persons Are Enduring Simple Spiritual Substances 155
7.3.1 Simple and Complex Views of Diachronic Personal Identity 155
7.3.2 Our Basic Awareness of Ourselves 156

- 7.3.3 Plantinga’s Replacement Argument 158
- 7.3.4 Normative, Teleological Rationality, Rational Deliberation, and the Simple Enduring Spiritual Self 160
  - 7.3.4.1 Reasoning Is Intrinsically Teleological 160
  - 7.3.4.2 Fulfillment Structures and Teleologically Arranged Action Plans 162
- 7.4 Defense of Premise (1): MAs Are Mereological Inconstant Objects 163
  - 7.4.1 An Ontological Assay of MAs 164
  - 7.4.2 Defeaters for the Complex View of PI 165
    - 7.4.2.1 The Circularity Problem 165
    - 7.4.2.2 Gradualism, Transitivity, and Sorites 166
    - 7.4.2.3 The Fission Problem 166
    - 7.4.2.4 Immanent Causation to the Rescue? 167
  - 7.4.3 Two Physicalist Alternatives to Human Persons as MAs 169
    - 7.4.3.1 We Are Atomic Simples 169
      - 7.4.3.1.1 Chisholm on Extended Physical Atoms or Simple Souls 169
      - 7.4.3.1.2 Mental Subjects Are Simple Souls: Assessing the Case 170
      - 7.4.3.1.3 The Ubiquitous Rejection of the EPA Position 171
    - 7.4.3.2 We Are Homeodynamic Systems 172
      - 7.4.3.2.1 Objections to the Systems Approach 173
        - 7.4.3.2.1.1 What We Seem to Know About Ourselves 174
        - 7.4.3.2.1.2 Avoid Spooky Entities at All Costs 174
        - 7.4.3.2.1.3 Complex Systems as Synchronically Unified Wholes 174
        - 7.4.3.2.1.4 Problems with Harmonizing Physicalism and Downward Causation (DC) 174
        - 7.4.3.2.1.5 Complex Systems as Enduring Continuants 178
        - 7.4.3.2.1.6 Versions of Staunch Hylomorphism Are Superior to Systems Theories of Organisms, Specifically, Human Persons 179

PART IV Updated and Novel Arguments from Modality and Libertarian Freedom 189

CHAPTER 8 Upgrading Modal Arguments for Substance Dualism 191
- 8.1 A Modal Argument for SD 191
  - 8.1.1 Statement of the Argument 191
  - 8.1.2 Clarification of the Argument 192
- 8.2 Defense of the Argument 193
  - 8.2.1 Contingent Physicalism and Premise (2) 193
    - 8.2.1.1 Merricks on Contingent Physicalism and the Modal Argument 193
      - 8.2.1.1.1 Objections to Merricks's Argument 194
    - 8.2.1.2 Contra Bailey's Contingent Physicalism 201
      - 8.2.1.2.1 Four Features of Bailey's Position 201
      - 8.2.1.2.2 Four Problems with Bailey's Position 203
      - 8.2.1.2.3 An Addition Response to Bailey's Via Negativa 211
  - 8.2.2 Clarification and Defense of Premise (3) 212
    - 8.2.2.1 Conceivability and Possibility 212
    - 8.2.2.2 Modal Epistemology, Rational Intuitions as Seemings and Adequate Intuitive Presentation 213
      - 8.2.2.2.1 Timothy O'Connor 213
      - 8.2.2.2.2 George Bealer 218
      - 8.2.2.2.3 Edmund Husserl 223
- 8.3 Five Objections Against Modal Arguments for SD 230
  - 8.3.1 The Modal Argument Is Circular 230
  - 8.3.2 The Modal Argument Establishes Merely a Duality of Concepts 231
  - 8.3.3 Physicalist and Dualism Intuitions Cancel Out Each Other 231
  - 8.3.4 Problems with Dualist Thought Experiments 232
  - 8.3.5 Kripke and the Confusion Between Epistemic and Metaphysical Possibility 233

CHAPTER 9 Staunch Libertarian Agency and the Simple, Enduring Soul 234
- 9.1 Contemporary Recognition of the Problem SLA Raises for Naturalism and Physicalism 235
- 9.2 The Argument and Its Basic Epistemic Justification 237
  - 9.2.1 The SLA Argument 237
  - 9.2.2 Our Fundamental Justification for Libertarian Freedom 237
- 9.3 SLA and Its Metaphysical Worldview Implications 240
  - 9.3.1 A Formal Characterization of SLA 241
  - 9.3.2 Six Features That Constitute SLA 241
    - 9.3.2.1 Free Agents Are Substances 241
    - 9.3.2.2 Free Agents Have Active Power 243
    - 9.3.2.3 Free Agents Are First Movers 243
    - 9.3.2.4 Free Agents Have Categorical Ability 245

9.3.2.5 Free Agents Act for the Sake of Teleological Ends 245
9.3.2.6 Free Agents Exhibit Top/Down Causation 246
9.4 Alternatives to SLA 249
9.4.1 Daniel Dennett and Free Will Irrealism 249
9.4.2 John Searle and Creative Compatibilism 250
9.4.2.1 Free Will Within Naturalist Constraints 250
9.4.2.2 A Solution to the Naturalist Problem of Free Will 251
9.4.2.3 A Critique of Searle’s Solution 254
9.4.3 Robert Kane and Faint-Hearted Libertarian Freedom 257
9.4.3.1 Kane and Naturalistic Constraints 257
9.4.3.2 Three Objections to Kane’s Position 259
9.4.3.2.1 Kane’s Denial of a Substantial Agent 259
9.4.3.2.2 Two Problems with Kane’s Account of Causality 259
9.4.3.2.3 *Kane*’s Rejection of Reasons as Teleological Ends 261
9.4.4 Kevin Timpe and Jonathan Jacobs: Minimalist Naturalism and Libertarianism 261
9.5 Causal Closure of the Physical and “Top-Down” Causation 263
9.5.1 Kim’s Supervenience Argument 263
9.5.2 Kim’s Crucial Background Assumption 263
9.5.3 The Major Difficulty: Emergence 264
9.5.3.1 No Causal Overdetermination 265
9.5.3.1.1 Ted Sider’s Counterargument and Our Response 266
9.5.3.1.2 Macro-Objects Are Aggregates 266
9.5.3.1.3 Adapting a Defeater from O’Connor and Churchill 267
9.6 Conclusion 271

PART V New and Neglected Responses to Common Defeaters Against Substance Dualism 273

CHAPTER 10 Important Frequently Raised Defeaters Against Substance Dualism 275
10.1 The Three Problems of Causal Interaction 276
10.1.1 Disambiguating “How Can Mental Entities Causally Interact with Physical Entities?” 276
10.1.2 Causal Interaction Violates the Conservation of Energy Principle 277
10.1.3 Causal Interaction Falls Prey to the Problem of Causal Pairing 280
10.2 Nine Neuroscientific Objections to SD 284
10.2.1 SD as a Soul-of-the-Gaps Argument 284
10.2.2 Neuroscience and the Explanatory Impotence of SD 285
10.2.3 Empirically Equivalent Theories and the Findings of Neuroscience 288

    - 10.2.4 Neuroscience and the Nature of the Central Issues in Philosophy of Mind 290
    - 10.2.5 Neuroscientific Methodology Relies on First-Person Reports 292
    - 10.2.6 Scientific Explanation, Necessitation, and the Failure of the Neuroscientific Correlations 293
    - 10.2.7 Neuroscientific Correlations and an Ad Hoc, Bloated Ontology 297
    - 10.2.8 The Neuroscientific Parsimony Objection 298
    - 10.2.9 Defeating and Turning the Tables on an SD Parsimony Argument 300
- 10.3 Postulating a Substantial Soul Is Explanatorily Impotent in Explaining the Occurrence of Conscious States 301
    - 10.3.1 SD and Explanatory Impotence 301
    - 10.3.2 SD and the Problematic Nature of Soul Stuff 302
    - 10.3.3 SD Responses to These Objections 303
    - 10.3.4 Responding to a Misguided Physicalist Defeater 303
    - 10.3.5 A Response to Bailey's Magical Mystery Tour 305

PART VI Charting a Future for Substance Dualism 309

CHAPTER 11 New Research Programs for 21st Century Substance Dualism 311

- 11.1 Conceptual Research Projects 312
    - 11.1.1 Charting the Logical Space of Mere Substance Dualism 312
        - 11.1.1.1 The Historical Argument 312
        - 11.1.1.2 The Substance Argument 314
        - 11.1.1.3 Varieties of Contemporary Substance Dualism 314
- 11.2 Methodological Research Projects 316
    - 11.2.1 Empirically Informed and Testable SD 316
    - 11.2.2 Ontologically Serious Philosophy of Mind 316
        - 11.2.2.1 Properties, Powers, and Substances 317
        - 11.2.2.2 Parts, Wholes, and Unity 318
        - 11.2.2.3 The Return of Teleology 318
    - 11.2.3 Preserving the Person/Subject 319
    - 11.2.4 Preserving the First-Person 319
    - 11.2.5 Consciousness-First and Analytic Phenomenology 320
- 11.3 Developmental SD Research Projects 322
    - 11.3.1 Charting the Logical Space of SD Embodiment 322
        - 11.3.1.1 Two-Dozen or so SD Embodiment Distinctions 322
        - 11.3.1.2 SD Embodiment and the Unity of the Person 324
        - 11.3.1.3 The Bodily Soul View 325
        - 11.3.1.4 Staunch Hylomorphic SD 326
    - 11.3.2 The Great Importance of Embodiment for SD 328

11.3.3 SD and Psychological Development 330
11.3.4 SD and the Nature of the Physical 331
11.4 Applied SD Research Projects 332
11.4.1 SD and New Waves in the Philosophy of Religion 332
11.4.2 SD and Epistemology 334
11.4.2.1 Internalism and Externalism 334
11.4.2.2 Social Epistemology 335
11.4.3 SD and Content Externalism 336
11.4.4 SD and Social Psychology 337
11.4.5 SD Healthcare and Mental Health 337
11.4.6 SD Approaches to Developing Technology 339
11.4.6.1 SD and Transhumanism 340
11.4.6.2 SD and Artificial Intelligence 340
11.4.6.3 SD and Virtual Reality 342
11.5 Conclusion 343

APPENDIX: The Soul, Mental Action, and the Conservation Laws 344
Mihretu P. Guta

Bibliography 361

Index 392

List of Figures

3.1 Genuine Wholes and Ordered Aggregates 49
3.2 Mereological Simplicity and Complexity 50
6.1 The Scattered Parts View 124
6.2 The Overlapping Parts View 131
7.1 Naturalist Mereological Hierarchy 149
7.2 Temporal Self-Awareness 156
8.1 Bailey’s Ontology 202
8.2 Husserl’s Epistemology of Defeasible Modal Knowledge 227
8.3 Husserl’s Ontology of *de re* Modality 229
9.1 Mental Causation 247
9.2 Naturalist Free Will 252
10.1 Wholly and Not-Wholly Material 300
10.2 Embodied and Non-Embodied 300
A.1 Lowe’s SD Mental Causation 352

Acknowledgments

We are grateful to the following people (listed in alphabetical order) who provided helpful feedback on various parts of an earlier draft of this book: Andrew Bailey, Todd Buras, William Lane Craig, Brian Cutter, C. Stephen Evans, Greg Ganssle, Robert Garcia, Stewart Goetz, Paul Gould, Mihretu Guta, Walter Hopp, Ross Inman, Greg Jesson, Sheki Lafanzio, Timothy O'Connor, Timothy Pickavance, Alexander Pruss, Joshua Rasmussen, Richard Swinburne, and Charles Taliaferro.

I (Moreland) want to give special thanks to my research assistants who went beyond the call of duty in helping me with research and editing: Jane Gropp, Eli Haitov, Hal Haller, Samantha Hawley, Jonathan Metcalf, Megan Reeve, and Mark Stanley.

I (Rickabaugh) want to give special thanks to the Martin Institute and the John Templeton Foundation for their very generous financial support, as well as Robert Garcia, Eff and Patty Martin, Steve Porter, Sam Kimbriel, James Catford, Michael Wear, and my diligent research assistants, Lucas Merritt, and Anna Ruth. Most importantly, to my wife, Laura, and our children, Alyssia, Sawyer, and Willow: your wellspring of love, support, play, and encouragement makes possible everything I do.

PART I

ONTOLOGICALLY SERIOUS PHILOSOPHY OF MIND

Contemporary and Tenacious Insights

CHAPTER 1

Substance Dualism in the 21st Century

1.1 THE RETURN OF SUBSTANCE DUALISM

1.1.1 The Great Revolt Against Dualism

At the end of the 19th century, substance dualism—roughly, the thesis that the human person is comprised of a substantial immaterial soul and a physical body—was widespread. Materialism was not a live option. As U.T. Place observed,

> [Ever] since the debate between Hobbes and Descartes ended in apparent victory for the latter, it was taken more or less for granted that whatever answer to the mind-body problem is true, materialism must be false.[1]

This sociological fact changed quickly, bringing about what William James described as "the evaporation of the definite soul-substance."[2] Arthur O. Lovejoy deemed the 20th century as "the Age of the Great Revolt against Dualism."[3] The inevitable defeat of substance

---

[1] Place, U. T. (2002). A pilgrim's progress? From mystical experience to biological consciousness. *Journal of Consciousness Studies*, *9*, 36. As quoted in Kind, A. (2019). The mind-body problem in 20th century philosophy. In A. Kind (Ed.), *Philosophy of mind in the twentieth and twenty-first centuries: A history of the philosophy of mind* (Vol. 6, p. 73). New York, NY: Routledge.

[2] James, W. (1904). Does "consciousness" exist? *The Journal of Philosophy, Psychology and Scientific Methods,* 1, 478.

[3] Lovejoy, A. O. (1930). The revolt against dualism: An inquiry concerning the existence of ideas (p. 1). La Salle, IL: Open Court Publishing.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

dualism was assumed to be a foregone conclusion. Gilbert Ryle had, in the words of Daniel Dennett, "danced quite a jig on the corpse of Cartesian dualism."[4]

Rebellious defenders of substance dualism remained (e.g., C. J. Ducasse (1881–1969), William McDougall (1871–1938), G. F. Stout (1860–1944), Karl Popper (1902–1994), John Eccles (1903–1997)). Physicalist William Lyon goes so far as to announce

> ... there is no doubt at all, that by the last decades of the nineteenth century and during the first [three] decades of the twentieth century, the new scientific psychologist's view of the correct way of finding out about minds, and implicitly of the nature of minds, was deeply Cartesian.[5]

Yet, despite a remarkable rise in apt defenders in the latter 20th century, novel and sophisticated defenses of substance dualism were largely ignored. In the words of D. M. Armstrong, dualism was seen as "curiously formal and empty."[6] Daniel Dennett incites the overall acceptance of "the dogmatic rule that dualism is to be avoided at all costs."[7]

Denunciations of substance dualism are often expressed quite dramatically. Here is a representative sample:

> "Few ideas are as unsupported, ridiculous and even downright harmful as that of the 'human soul'."[8]
>
> ... dualism is akin to explaining lightning in terms of Thor's anger, and hence is fundamentally primitive and pre-scientific.[9]
>
> ... the concept of a non-physical soul looks increasingly like an outdated theoretical curiosity.[10]
>
> To defend dualism is to close one's eyes to the enormous and ongoing scientific developments in the mind and brain sciences in the past few centuries—and to wallow in mystery.[11]
>
> "Immaterial mind" or "soul," like "élan vital," "elf," or "chupacabras," are ghostly expressions that come from mistaken frameworks or conceptions and do not refer to anything.[12]
>
> There is no spirit-driven life force, no throbbing, heaving, pullulating, protoplasmic, mystic jelly. Life is just bytes and bytes and bytes of digital information.[13]

---

[4] Dennett, D. (1987). *The intentional stance* (p. 214). Cambridge, MA: MIT Press. Dennett is referring to the arguments in Ryle's 1949 book, *The Concept of Mind*.

[5] Lyons, W. (2001). *Matters of the mind* (p. 14). New York: Routledge; cf. 1–36, especially 13–24. Brackets ours.

[6] Armstrong, D. M. (1968). *A materialist theory of mind* (p. 23). London: Routledge and Kegan Paul.

[7] Dennett, D. (1991). *Consciousness explained* (p. 37). Boston, MA: Little, Brown and Co.; his italics.

[8] Barash, D. P. (2023). Stuck with the soul. *Aeon Essays*. (March 20, 2023). Retrieved from https://aeon.co/essays/why-are-most-of-us-stuck-with-a-belief-in-the-soul.

[9] Braddon-Mitchel, D., & Jackson, F. (1996). *Philosophy of mind and cognition* (p. 8). Oxford: Blackwell.

[10] Churchland, P. (2002). *Brain wise: Studies in neurophilosophy* (p. 173). Cambridge, MA: MIT Press.

[11] Schouten, M. K. (2001). Theism, dualism, and the scientific image of humanity. *Zygon*, *36*(4): 696.

[12] Sabatés, M. H. (2005). Reductionism in the philosophy of mind. *Encyclopedia of Philosophy*. Retrieved from http://www.encyclopedia.com/humanities/encyclopedias-almanacs-transcripts-and-maps/reductionism-philosophy-mind.

[13] Dawkins, R. (1995). *River out of Eden: A Darwinian view of life* (p. 18). New York, NY: Basic Books.

Notice that these quotations employ mere rhetorical devices. John Searle has pointed out at least four rhetorical devices used in the philosophy of mind to either ridicule opposing views or hide the implausibility of one's view.[14] Searle's observations apply equally to so many dismissals of substance dualism.

1. The use of technical jargon to cover up the implausibility of one's view.[15]
2. The stylistic use of prose to evade clearly stating one's argument.[16]
3. Naming an opposing view (e.g., "the Cartesian intuition"), to which one then denies by name and not by content.[17]
4. The covering up of one's troubling view by means of an analogy with a scientific discovery of the past that also seemed silly or ignorant.[18]

To Searle's list, we add a fifth device.

5. The practice of associating substance (and, sometimes, property) dualists with accepting the existence of spooks or spooky stuff, ghosts, or other paranormal entities is best left to the pages of *The National Enquirer*. These labels function as a way of associating dualism with astrology, the occult, or other superstitious views contrasting with physicalism, which is the clear implication of tough-minded progress in the hard sciences.[19]

One would hope that such unabashed ridicule would reflect the great strength of anti-dualist arguments. But, as we will demonstrate in the chapters to come, that hope is questionable if not outright ungrounded.

1.1.2 The Resurgence of Substance Dualism

Fast forward to the late 20th and early 21st centuries. Things are changing. We are witnessing a resurgence of substance dualism. Physicalism remains the dominant view in the philosophy of mind. Yet, it is no exaggeration to say that substance dualism is undergoing an unforeseen revival and is poised to make a strong return in the 21st century. These reasons are not mere sociological reflections but expressions of the rapid growth in sophisticated new work on substance dualism. A large number of recently published book-length defenses of substance dualism evidence this, along with collections of new pro-substance dualism papers. The number of journal articles defending substance dualism is too long to list.[20] Conspicuous by their absence are substantive engagements with these works.

---

[14] Searle, J. *Rediscovery of the mind* (p. 9). Boston, MA: MIT Press.

[15] Ibid., 9.

[16] Ibid., 4.

[17] Ibid., 4–5.

[18] Ibid., 5.

[19] In addition to Gilbert Ryle's (infamous) statement about "the ghost in the machine," see Kim, J. (2018). Against cartesian dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 154–167). Oxford: Wiley-Blackwell. The quote is on 153; Churchland, P. (2013). *Matter, and consciousness* (3rd revised ed., pp. 12–16). Cambridge, MA: MIT Press; Murphy, N. (2007). Reductionism: How did we fall into it and can we emerge from it? In N. Murphy & W. R. Stoeger (Eds.), *Evolution and emergence* (pp. 19–39). Oxford: Oxford University Press; Steward, H. (2012). *A metaphysics for freedom* (p. 228). Oxford: Oxford University Press.

[20] An exhaustive bibliography is available at www.brandonrickabaugh.com.

1.2 THE CASE AGAINST SUBSTANCE DUALISM WEAKENS

1.2.1 Admissions to the Weakness of Standard Objections to Substance Dualism

In addition to the resurgence of substance dualism, some non-dualists argue that standard objections to substance dualism are less triumphant than previously thought. After working through standard arguments against substance dualism, José Gusmão Rodrigues, himself a Russellian monist, concludes,

> A significant part of current literature on the mind-body problem, personal identity and other general metaphysical questions seems to me rife with anti-dualistic assumptions without good arguments to back it [*sic*] up. Such an attitude is dogmatic and contrary to good philosophical spirit.[21]

Likewise, materialism William Lycan admits,

> Being a philosopher, of course I would like to think that my stance is rational, held not just instinctively and scientistically and in the main-stream but because the arguments do indeed favor materialism over dualism. But I do not think that, though I used to. My position may be rational, broadly speaking, but not because the arguments favor it: Though the arguments for dualism do (indeed) fail, so do the arguments for materialism. And the standard objections to dualism are not very convincing; if one really manages to be a dualist in the first place, one should not be much impressed by them.[22]

In Chapter 10, we will argue in detail that the standard objections to substance dualism fail in ways even staunch anti-dualists must acknowledge.

1.2.2 Admitting the Faith of Anti-Dualism

Another striking phenomenon is the growing recognition that eschewing substance dualism is essentially the expression of faith in naturalism, faith in what Karl Popper referred to as promissory materialism: that in the not-too-distant future, science will provide an in-principle explanation of all phenomena in the world, including consciousness.[23]

David Chalmers observes, “The main residual motivation to reject substance dualism may simply lie in the term’s negative connotations.”[24] If Chalmers is correct, then the rejection of substance dualism is most likely driven primarily by mere ideological allegiance—by faith—rather than rigorous argument. John Searle expresses this posture well.

---

[21] Rodrigues, J. G. (2014). There are no good objections to substance dualism. *Philosophy*, *89*(2), 221–222.

[22] Lycan, W. (2009). Giving dualism its due. *Australasian Journal of Philosophy*, *87*(4), 551.

[23] Cf. Popper, K. R., & Eccles, J. C. (1985). *The self and its brain* (pp. 96–98). New York, NY: Springer-Verlag.

[24] Chalmers, D. (1996). *The conscious mind* (p. 170). Oxford: Oxford University Press.

> [We] have a terror of falling into Cartesian dualism. The bankruptcy of the Cartesian tradition and the absurdity of supposing that there are two kinds of substances or properties in the world, "mental" and "physical," is so threatening to us and has had such a sordid history that we are reluctant to concede anything that might smack of Cartesianism.[25]

As a result, substance dualism has not been tried and found wanting so much as it has been judged unacceptable and left untried.

1.2.3 Proposals of Naturalistic Substance Dualism

Indeed, the presumption of naturalism and atheism fuels most rejections of substance dualism. Why else would the guilt by religious association objection be relevant? This habit is not well thought out; some naturalists acknowledge as much. As Michael Martin, perhaps one of the leading atheists of the past 100 years, states, "atheism is not materialism ... atheism is compatible with various forms of mind-body dualism."[26] Graham Oppy makes the same observation.[27] Atheist Evan Fales observes,

> A naturalist who is not sanguine about the reducibility of the mental to the physical has several options. These include token-token identity theories such as functionalism, property dualism, and—more daringly—substance dualism.[28]

Naturalism, according to Fales, plausibly commits one to reject the existence of disembodied minds. But substance dualism does not entail disembodied minds, nor is naturalism incompatible with substance dualism.[29] Likewise, naturalist Eric Steinhart explains,

> While many naturalists dogmatically reject such mind-body dualism, progressive naturalists are open to it. They ask only for a research program which formally defines the immaterial mind and its linkage with the brain. After all, since mathematical structures need not be material, the immateriality of the mind cannot block its formal definition.[30]

As examples, Steinhart refers to Ray Kurzweil and Hillary Putnam's concession that computers are made of purely mental stuff.[31] While some dissenters are committed theists, nearly all naturalists and atheists reject substance dualism. However, to the surprise of many, some atheists and agnostics defend substance dualism (e.g., C. D. Broad, Karl Popper, and Michael Huemer).[32] Still, as we will see in Chapter 2, there are good reasons to think that a naturalist should eschew substance and property dualism.

---

[25] Searle, J. (1992). *The rediscovery of the mind* (p. 13). Cambridge, MA.

[26] Fernandes, P., & Martin, M. (2000). *Theism vs. atheism: The internet debate* (p. 19). IBD Press.

[27] Oppy, G. (2013). *The best argument against god* (p. 55). New York, NY: Palgrave Macmillan.

[28] Fales, E. (2007). Naturalism and physicalism. In M. Martin (Ed.), *The Cambridge companion to atheism* (p. 127). New York, NY: Cambridge University Press.

[29] Ibid., 127–128.

[30] Steinhart, E. (2019). Naturalism. In G. Oppy (Ed.), *A companion to atheism and philosophy* (p. 159). New York, NY: Wiley-Blackwell.

[31] Steinhart cites Richards, J. W. (Ed.). (2002). *Are we spiritual machines? Ray Kurzweil vs. the critiques of strong AI*. Seattle, WA: Discovery Institute; Putnam, H. (1975). Philosophy and our mental life. In *Mind, language, and reality*. Cambridge: Cambridge University Press.

1.3 THE REVIVAL OF SUBSTANCE DUALIST INTUITIONS

Another significant change is the recognition of intuitions against standard physicalism that have historically been used to defend property and substance dualism. Although not exhaustive, we note four.

1.3.1 The Return to the Self and Subject of Consciousness

A mark of 20th-century philosophy was the rejection of the self and attention to the subject of consciousness. Most leading philosophers of mind outright express the first. The second, if not explicit, is indeed implicit in the widespread practice of analyzing mental states without metaphysical considerations of the subject of mental states. Standard discussions of consciousness, for example, predicate phenomenal consciousness or qualia to mental states. One rarely hears of subjects—the bearers of mental states, as the thing that is conscious. Substance dualists have not followed suit but have kept the self and subject of consciousness as a necessary component of the analysis of consciousness. In Chapter 5, we argue why this must be the case for any successful analysis of consciousness. Here we note that the self and subject of consciousness are returning to the fore of consciousness studies. And resistance to taking the self and subject ontologically seriously is resisted by its obvious connection to substance dualism. Yet, the intuition of the reality of the self and subject is too difficult for many to resist.

1.3.2 The Meta-Problem of Consciousness and Cognitive Science

There is also a growing recognition that dualist intuitions that consciousness is not physical and the intuition that the subject of consciousness is an immaterial substance must be taken seriously. Once rejected as the product of ignorant folk psychology, the ubiquity and strength of pro-dualist intuitions are being recognized as essential data to be explained. The substance dualist has understood this all along. We will say more about this at the end of the next section.

1.3.3 The Revival of Aristotelianism

The growing revival of Aristotelianism during the last two decades, especially in its recognition of substances and their indispensability in solving various metaphysical problems, is a significant step away from standard physicalism and toward substance dualism. As we will see in Chapter 7, this revival has two fundamental versions—weak and strong hylomorphism.

---

[32] See, e.g., Broad, C. D. (1959). Autobiography. In P. A. Schilpp (Ed.), *The philosophy of C.D. Broad* (pp. 3–68). New York: Tudor; see 43–44; Huemer, M. (2018). *Paradox lost: Logical solutions to ten puzzles of philosophy* (p. 235). Switzerland: Palgrave Macmillan; and Zerin, E. (1998). Karl Popper on God: The lost interview. *Skeptic*, *6*(2), 47–48.

The former reduces the substantial form (or essence) to a relational structure constituted by external relations whose relata are separable parts. This version comes perilously close to some form of vague naturalistic animalism.

But strong hylomorphism entails a more classic understanding of substances as essentially characterized particulars that, synchronically, have properties and inseparable parts as constituents and, diachronically, retain Leibnizian identity through teleological processes aimed toward an end. Here, the substantial form serves (at least) three metaphysical functions: (i) It unifies the substance synchronically; (ii) it grounds the substance's diachronic identity; (iii) it is the teleological guide of the lawlike sequences of development for the sake of achieving its proper end.

When we talk of a revival of Aristotelianism, we refer primarily to the resurgence of strong hylomorphism in analyzing the nature of living organisms, though that revival is much broader.[33]

1.3.4 Broad Worldview Considerations

Over the last fifty years, there has been a ubiquitous resurgence of theism, predominantly Christian theism, across the philosophical disciplines. The result? Naturalism is no longer the only game in town. With the loss of its hegemony, broader worldview implications, particularly those involving naturalism versus theism, were back on the table. In the next chapter, we will apply Thomas Nagel's recent insight that in the current setting, one must keep broad worldview issues in mind while working in philosophy of mind and conversely.

Given the ascendency of theism, proponents articulated several claims about the advances in theistic scholarship and the explanatory resources of theism vis-à-vis naturalism in the philosophy of mind, e.g., precise clarification of the intelligibility of a Spiritual Substance, additional insights on various interaction problems, the superior explanatory power of theism regarding the origin of consciousness and several features of human persons. Human persons fit naturally in a theistic worldview but remain recalcitrant for naturalism. We will take up a number of these in the chapters to follow.

1.4 THE CURRENT TURNING AWAY FROM STANDARD PHYSICALISM

Due to the persistent failures to solve the hard problem of consciousness, a growing number of philosophers are turning away from standard physicalism (a.k.a. strong physicalism). In 2005, David Chalmers compiled a list of at least 30 noted contemporary anti-physicalists.[34] Of course, there are materialists on this list (e.g., Trenton Merricks), but materialists about human persons needn't be physicalists.

---

[33] See Oderberg, D. (2007). *Real essentialism*. New York, NY: Routledge; Feser, E. (2014). *Scholastic metaphysics*. Heusenstamm, Germany: Editiones Scholasticae; Ellis, B. (2002). *The philosophy of nature: A guide to the new essentialism*. Bucks, Great Britain: Acumen Press; Lowe, E. J. (2006). *The four-category ontology*. Oxford: Oxford University Press; Brown, C. M. (2005). *Aquinas and the ship of Theseus*. London: Continuum; Simpson, W. M. R., Koons, R. C., & Nicholas, J. (Eds.). (2018). *Neo-Aristotelian perspectives on contemporary science*. New York, NY: Routledge. For additional resources, see Chapter 7, notes 90, 92, 101–102, and 105–112.

1.4.1 The Fundamentality of Consciousness

Among physicalist dissenters are those now positing phenomenal consciousness as a fundamental feature of reality (e.g., panpsychism and Russellian monism).[35] This is significant for at least three reasons. First, as Charles Taliaferro rightly points out, the substance dualists and panpsychists both (i) recognize the irreducible reality of consciousness; (ii) resist reductive accounts of consciousness that at least implicitly result in eliminativism; and (iii) tend to reject accounts that take consciousness as both irreducible and emergent from brute matter entirely bereft of mental elements.[36] Just as some have argued that property dualism leads to substance dualism,[37] the shift toward panpsychism is likewise significant.

1.4.2 The Return of Panpsychism

Second, the shift toward panpsychism rides on certain pro-dualist intuitions. There is a growing recognition of what substance dualists have always found intuitive, that the mental is so different from the physical that consciousness is inexplicable, even ontologically disallowed, within the limitations of physicalism. Hence, a proposal more radical than physicalism is needed. In light of this, Chalmers observes,

> I think that substance dualism (in its epiphenomenalist and interactionist forms) and Russellian monism (in its panpsychist and panprotopsychist forms) are the two serious contenders in the metaphysics of consciousness, at least once one has given up on standard physicalism. (I divide my own credence fairly equally between them).[38]

---

[34] His list included Joseph Almog, Torin Alter, George Bealer, Laurence BonJour, Paul Boghossian, Tyler Burge, Tim Crane, John Foster, Brie Gertler, George Graham, W.D. Hart, Ted Honderich, Terry Horgan, Steven Horst, Jaegwon Kim, Saul Kripke, Harold Langsam, E.J. Lowe, Kirk Ludwig, Trenton Merricks, Martine Nida-Rümelin, Adam Pautz, David Pitt, Alvin Plantinga, Howard Robinson, William Robinson, Gregg Rosenberg, A.D. Smith, Richard Swinburne, and Stephen White. See Chalmers, D. (2005, September 26). Fragments of consciousness: A weblog by David Chalmers. https://fragments.consc.net/djc/2005/09/jaegwon_kim_com.html.

[35] See, e.g., Brüntrup, G., & Jaskolla, L. (Eds.). (2017). *Panpsychism: Contemporary perspectives*. Oxford: Oxford University Press; and Alter, T., & Nagasawa, Y. (Eds.). (2015). *Consciousness in the physical world: Perspectives on Russellian monism*. Oxford: Oxford University Press.

[36] Taliaferro, C. (2017). Dualism and panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (pp. 383–384). Oxford: Oxford University Press.

[37] See, e.g., Lycan, W. G. (2013). Is property dualism better off than substance dualism? *Philosophical Studies*, *164*(2), 533–542; Schneider, S. (2012). Why property dualists must reject substance physicalism. *Philosophical Studies*, *157*(1), 61–76; Zimmerman, D. (2010). From property dualism to substance dualism. *Aristotelian Society Supplementary*, *84*(1), 119–150; and Francescotti, R. (2001). Property dualism without substance dualism? *Philosophical Papers*, *30*(2), 93–116.

[38] Chalmers, D. (2017). Panpsychism and panprotopsychism. In Brüntrup & Jaskolla (pp. 39–40).

That a leading philosopher of mind would consider substance dualism among the two serious contenders would be unthinkable in the early 20th century. And as we will argue in a moment, specific panpsychist commitments make resisting substance dualism very difficult.[39]

1.4.3 Consciousness-First Philosophy of Mind

Likewise, consider the developing consciousness-first approach to the philosophy of mind. Harold Langsam draws attention to how physicalism proceeds from commitments rejected by dualists.[40] Whereas physicalists privilege third-person data regarding the non-phenomenal, structural, and causal features associated with consciousness; dualists favor first-person data focused on the phenomenological qualities of consciousness.[41] Given this disparity of starting points, Langsam urges the dualist to ignore the physicalist's framing of the debate and develop a positive project showing how the dualist conception of the mind is indispensable.[42]

Thus, the consciousness-first approach rejects reductive explanations and reductive ontologies of consciousness.[43] Instead, consciousness is taken as a starting point from which to explain the nature and possibility of intentionality, mental causation, rationality, and normativity.[44] Central to the consciousness-first approach is a commitment to the foundational role that knowledge by acquaintance plays in philosophical analysis.[45] Again, we have an example of the recent embrace of pro-dualist intuitions and methodology. Chapter 11 will chart a research program for consciousness-first substance dualism.

1.4.4 Phenomenal Transparency and the Opacity of the Physical

Third, in conjunction with rejecting standard physicalism, these pro-dualist intuitions make resisting introspective arguments for substance dualism challenging. While physicalists typically deny such knowledge, Russellian monism and Russellian panpsychism embrace the following thesis:

---

[39] For an interesting argument that panpsychism can't avoid substance dualism, see Weir, R. S. (2021). Can a post-Galilean science of consciousness avoid substance dualism? *Journal of Consciousness Studies*, *28*(9–10), 212–228.

[40] Langsam, H. (2001). Strategy for dualists. *Metaphilosophy*, *32*(4), 396.

[41] This isn't to further the misconception that, unlike physicalists, dualists do not seek compatibility with any relevant empirical science. For an excellent criticism of this misconception, see Gertler, B. (2020). Dualism: How epistemic issues drive debates about the ontology of consciousness. In U. Kriegel (Ed.), *The Oxford handbook of the philosophy of consciousness* (pp. 277–300). Oxford: Oxford University Press.

[42] According to Langsam, this strategy works only for property dualism and not substance dualism, although he offers no clear reason for why we should accept this.

[43] Adam Pautz, for example, argues that reductive accounts of consciousness cannot accommodate the various ways in which consciousness is significant (dissimilarity-grounding, reason-grounding, and thought-grounding), although non-reductive accounts of consciousness can. See Pautz, A. (2017). The significance argument for the irreducibility of consciousness. *Philosophical Perspectives*, *31*(1), 349–407.

[44] See, e.g., Langsam, H. (2002). Consciousness, experience, and justification. *Canadian Journal of Philosophy*, *32*(1), 1–28; Pautz, A. (2013). Does phenomenology ground mental content? In U. Kriegel (Ed.), *Phenomenal intentionality: New essays* (pp. 194–234). Oxford: Oxford University Press.

[45] For recent work on the acquaintance approach to consciousness studies, see, e.g., Smithies, D. (2019). *The epistemic significance of consciousness*. Oxford: Oxford University Press; Gertler, B. (2012). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 93–127). Oxford: Oxford University.

> *Phenomenal Transparency*: Phenomenal concepts reveal the nature of the conscious states to which they refer.

When a subject attends to her nonconceptual acquaintance of her conscious states, she is provided with a conceptual understanding of those conscious states.[46] I know what pain is by attending to my direct first-person acquaintance of my experience of being in pain, in virtue of which I can form a conception of pain that is phenomenally transparent, revealing the nature of pain.

Various arguments against physicalism (e.g., the conceivability, knowledge, and explanatory-gap arguments) depend on a thesis like phenomenal transparency, according to which we have epistemic access to the nature of consciousness.[47] Such an anti-physicalism argument, says Goff, "hangs and falls on Phenomenal Transparency," and he spends three chapters explaining why.[48]

Here is the problem. By embracing Phenomenal Transparency, the panpsychist must take the dualist intuition seriously that one's conception of oneself as a subject of consciousness is likewise phenomenally transparent. Moreover, they cannot avoid this and similar arguments for substance dualism by following the standard physicalist rejection of knowledge by acquaintance with one's conscious states.[49] Additionally, we will see in Chapter 10 that the rejection of Phenomenal Transparency, and self-knowledge by acquaintance, has unknowingly led many physicalists to promote a grotesque strawman according to which substance dualism commits one to soul-stuff that can only be characterized by way of the *via negative*.

Embracing a further pro-dualist thesis, some recent panpsychists and non-substance dualists defend non-deflationary accounts of subjecthood. Goff, for example, argues that (direct) Phenomenal Transparency entails a non-deflationary account of subjecthood because it allows one to grasp the essence of any phenomenal property and, in doing so, the essence of subjecthood. This entails that subjects cannot be explained away.[50]

To be clear, we are not defending introspection arguments for substance dualism. That will be ventured in Chapter 4. Instead, we are merely pointing out that the current turn toward panpsychism evidences a new acceptance of pro-dualist intuitions (e.g., direct acquaintance with one's mental states and the irreducibility of phenomenal consciousness), which in turn increases the probability of substance dualism.

Lastly, consider a thesis of Russellian panpsychism.

> *Opacity of the Physical*: While we know the dispositional properties of the physical, we lack knowledge of its qualitative phenomenal properties.

---

[46] Goff, P. (2017). *Consciousness and fundamental reality* (p. 115). Oxford: Oxford University Press. See Chapter 9 for more on this.

[47] Chalmers, D. (2010).The two-dimensional argument against materialism. In D. Chalmers, *The character of consciousness* (p. 141). Oxford: Oxford University Press.

[48] Goff, 17, and chapters 3–5.

[49] See, e.g., Churchland, P. (2013). *Matter and consciousness* (3rd ed., pp. 24–25). Cambridge, MA: MIT Press.

[50] Goff, 178–181, 210–211.

Our knowledge of the physical is so incomplete that we have failed to understand that phenomenal consciousness is an aspect of the physical.

Here comes a problem. If the physical is so opaque as to conceal its phenomenal properties, the standard interaction problem loses its teeth. That form of the interaction problem is motivated by the claim that the mental is too dissimilar from the physical to allow any causal interaction between them. The Russellian panpsychist must hold that we know enough about the physical to motivate the interaction problem while also maintaining that the physical is opaque enough to suppose that matter, in some sense, has phenomenal properties. However, what in-principle reason could be given for thinking that the physical is not so opaque as to blind us to the fact that it cannot causally interact with the mental? Consequently, the recent turn toward Russellian panpsychism increases the probability of substance dualism by significantly weakening a major version of the interaction problem—more on this in Chapter 10.

In summary, substance dualism is returning. The recent move toward panpsychism and Russellian monism signals a growing embrace of pro-dualist intuitions and methodologies, raising the probability of substance dualism. And there is an increasing recognition that we should expect the nature of consciousness to be weird or far stranger than was anticipated in the 20th century. Commenting on the rejection of panpsychism as a crazy view, David Chalmers observes, "While the view is counterintuitive to some, there is good reason to think that any view of consciousness must embrace some counterintuitive conclusions."[51] That substance dualism strikes some as odd shouldn't mean much of anything. The 20th-century's *certain-defeat-of-dualism* narrative can no longer be presumed. In fact, substance dualism in the 21st century seems quite promising.

1.5 MERE SUBSTANCE DUALISM

1.5.1 Delineating the View

Our goal in this book is to defend substance dualism in its broadest form, what we call *Mere Substance Dualism*. In general, this minimal thesis is all that most arguments for substance dualism entail and is the starting point for any specific views. Substance dualism, explains Dean Zimmerman, is the view that,

> for every person who thinks or has experiences, there is a thing—a soul or spiritual substance—that lacks many or most of the physical properties characteristic of non-thinking material objects like rocks and trees; and that this soul is essential to the person, and in one way or another responsible for the person's mental life.[52]

Similarly, William Hart explains substance dualism as the thesis according to which "... each person's mind is at least not identical with his body, so these are two different things."[53]

---

[51] Chalmers, Panpsychism and panprotopsychism, 19.

[52] Zimmerman, D. (2010). From property dualism to substance dualism. *Aristotelian Society Supplementary*, *84*(1), 119–120. See also Zimmerman, D. (2005). Dualism in the philosophy of mind. In D. M. Borchert (Ed.), *Encyclopedia of philosophy* (2nd ed., Vol. 3, p. 113). Macmillan.

[53] Hart, W. D. (1994). Dualism. In S. Guttenplan (Ed.), *A companion to the philosophy of mind* (p. 265). Oxford: Blackwell Publishers.

This common understanding of substance dualism emphasizes the nature of the self or soul as an immaterial substance and leaves the nature of the body an open question.[54]

Substance dualism requires the following theses:

(a) The human person (*sans* disembodiment) is comprised in one way or another of a soul and a body.[55]
(b) The soul is an immaterial substance.[56]
(c) The body is—or at least partially is—a physical entity or an aggregate of physical entities.
(d) In principle, the human person can exist without his/her body.

There are further commitments of substance dualism we should consider.

(e) The mental life of a human person is possessed and unified by his/her soul.

Importantly, (e) does not entail that the soul is fully responsible for one's mental life but allows the brain and central nervous system (if not the whole body) to play a role in one's mental life. All that (e) requires is that the soul is necessary for the mental life of a human person.

(f) The soul is fundamental.

Issues of fundamentality are complex. Broadly speaking, we take (f) to mean that the soul is not reducible (ontologically, linguistically, causally, or explanatorily) to anything else.

Taking (a) through (f) together, we get the following thesis:

> *Mere Substance Dualism* (SD): The human person (i) is comprised of a soul (a fundamental, immaterial/spiritual substance) and a physical body, (ii) capable of existing without a body, but not without his/her soul, and (iii) the mental life of which is possessed and unified by his/her soul.

This is the thesis we will defend throughout this book, although, at times, we will include details from specific substance dualist views. We will refer to *Mere Substance Dualism* as substance dualism or SD for brevity.

In conclusion, many arguments against naturalism and physicalism and arguments for SD are too often neglected in the philosophy of mind. And when these arguments are presented, it is often the case that the presentation is too brief to examine the depth of those arguments. Among other things, we hope to remedy these problems in the following chapters.

---

[54] Those who state SD similarly include Howard Robinson, Geoffrey Madell, and N. M. Nathan. Cf. Robinson, H. Substance dualism and its rationale. In R. Swinburne (Ed.), *Free will and modern science* (p. 162). Oxford: Oxford University Press; Madell, G. (2010). The road to substance dualism. *Royal Institute of Philosophy Supplement*, *67*, 56; Nathan, N. M. L. (2011). Substance dualism fortified. *Philosophy*, *86*(02), 201.

[55] We use "comprised" to remain neutral regarding issues of composition or constitution. The substance dualist, *qua* substance dualist, needn't commit to such specifics. Moreover, our way of putting it leaves open the possibility that the person is identical to that person's soul and, in the ordinary scheme of things, has a body that includes physical entities.

[56] Staunch hylomorphists view substantial form as an abstract particular and call it a soul that departs at death and, somehow, sustains absolute personal identity. Here, the soul *qua* particularized substantial form is substance-like, i.e., it is an incomplete substance that requires a body to form a genuine substance. We include these views within our use of Mere Substance Dualism.

1.5.2 The Undeniable Recalcitrance of Substance Dualism

Substance dualism has withstood centuries of objections. It is a minority view with a resurgence underway. SD is almost always dismissed as out of fashion. However, without the dubious assumption that current ideological popularity is a defeater for argument, this dismissal is telling. But suppose that ideological popularity provides some vital evidence against a view. Now the story changes dramatically. SD enjoys greater evidential support than naturalist views, as SD is arguably the most popular view in the entire history of philosophy. And appealing to current ideological fashion's popularity seems arbitrary, historically chauvinistic, or naturalistically presumptuous.

In the following chapters, we will lay out a detailed, sometimes new, case for SD. But as noted above, we must keep an eye on worldview issues when doing philosophy of mind. Thus, to those issues, we now turn.

CHAPTER 2

How Staunch Naturalism Shapes the Dialectic in Philosophy of Mind

One of the central purposes of this book is to show how worldviews influence work in philosophy of mind and conversely. This is predominantly true regarding the worldview(s) of philosophical or scientific naturalism.[1] So, before we present arguments for and against SD in Chapters 4–10, we will focus our attention on worldview issues, issues that are often neglected and not made explicit in the literature in philosophy of mind.

2.1 WORLDVIEWS, SCIENTIFIC NATURALISM, AND THE STANDARD PICTURE

2.1.1 Philosophy of Mind and Worldview Implications

As Thomas Nagel reminded us, topics in the philosophy of mind—and action theory—are not local; they inform our understanding of the entire cosmos and its history and conversely. Perfect examples of this are the numerous anti-dualists quotations discussed in Chapter 1. If irreducible consciousness, substantial souls/selves, libertarian agency, and objective reason exist, we must ask what sort of reality could and did give rise to them.

---

[1] For a detailed examination of this claim, see Moreland, J. P. (2009). *The recalcitrant imago dei: Human persons and the failure of naturalism*. London: SCM Press.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

And, says Nagel, theism becomes a major player in that discussion. Thus, one must keep an eye on the worldview implications of a position on a topic in philosophy of mind and on how one's worldview informs one's positions in philosophy of mind.[2]

To be sure, doing things like offering a precise version of causal-role functionalism, along with proffering arguments for it, generally do not involve explicit reference to theism, naturalism, and the like. But when it comes to the methodology employed in philosophy of mind (e.g., the value of first-person "knowledge" relative to third-person claims, serious vs. shopping-list metaphysics, the relationship between metaphysics and natural science), the sought-after goals, the weighing and weighting of arguments, and background considerations derived from broad conceptual schemes, especially from worldviews, are crucially important.

2.1.2 Two Versions of a Naturalistic Worldview

In light of this, we should clarify what we mean by some important terms to be employed regularly in what follows, especially staunch/faint-hearted naturalism and strong/weak physicalism.[3] *Staunch naturalism* is the view that the cosmos and all entities within it—e.g., particulars, properties, relations, events, processes—are and only are physical. *Strong physicalism* adopts the same ontology concerning consciousness, the human person, and other entities studied in the philosophy of mind. While he does not use the term, Galen Strawson's understanding of standard physicalism and his claim that it is the ubiquitously adopted view underscores our assertions about *staunch naturalism*: "I take [standard] physicalism to be the view that every real, concrete [i.e., spatio-temporally located] phenomenon in the universe is … physical."[4] According to Strawson, standard physicalism entails (NE): "Physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential."[5]

Staunch naturalists/strong physicalists eschew *sui generis,* irreducible, ineliminable emergent properties.[6] *Faint-hearted naturalism* and *weak physicalism* differ only in that they embrace emergent entities, usually properties, but in a few cases, processes, systems, and even substances. Examples of staunch naturalism's influence in philosophy in general, and philosophy of mind in particular, are ubiquitous. Here are some examples.

---

[2] Nagel, T. (2012). *Mind and cosmos*. Oxford: Oxford University Press.

[3] Koons, R. (2014). Staunch vs. fainthearted hylomorphism: Towards an Aristotelian account of composition. *Res Philosophica*, *91*(2), 151–177.

[4] Strawson, G. (2006). Realistic monism: Why physicalism entails panpsychism. In A. Freeman (Ed.), *Consciousness and its place in nature* (pp. 3–31). Exeter, UK: Imprint Academic. The quote is from page 3. Strawson's idiosyncratic use of "physicalism" includes irreducible experiential states as a kind of physical state, and he distinguishes his view from "physicalism," the faith commitment that the nature or essence of all concrete reality can in principle be captured in the terms of physics. The definition of "physicalism" in the quote is not his own panpsychist one but is more akin to physicalism and what we mean by staunch naturalism. Brackets ours.

[5] Ibid., 11.

[6] For our purposes, we set aside the distinction between Global Naturalists (who deny the existence of abstract objects) and Local Naturalists (who accept abstract objects, usually with the proviso that the only abstract objects that are exemplified must be physical properties or relations). For a good treatment of this distinction put in different terms and including a case that naturalists ought to be Global, see Grossmann, R. (1992). *The existence of the world* (pp. 1–46). London: Routledge; Moreland, J. P. (2000). Naturalism and the ontological status of properties. In W. L. Craig & J. P. Moreland (Eds.), *Naturalism: A critical analysis* (pp. 67–109). London: Routledge.

2.1.3 Naturalism and the Standard Picture

Daniel Stoljar rightly identifies "the Standard Picture" of what philosophy is supposed to be about for most contemporary philosophers.[7] According to Stoljar, we all seem to be presented with a commonsense, Manifest Image of reality: People are substantial conscious selves who experience pains and see objects that are actually colored, who act intentionally and with libertarian agency, and for whom normative reasons affect what they do or don't do. But given the widely accepted worldview of (staunch) naturalism and the strong physicalism it underwrites, the Standard Picture constitutes the aim and informs the views and methodology of the discipline of philosophy as follows:

1. [Strong] physicalism is true.
2. [Strong] physicalism summarized the picture of the world implicit in the natural sciences.
3. It is most rational to believe the picture of the world implicit in the natural sciences.
4. [Strong] physicalism is *prima facie* in conflict with many everyday life presuppositions (the Manifest Image).
5. The way to resolve these conflicts is to propose views about how to understand these presuppositions, so they are compatible. This will involve some version of reduction or elimination of the problematic entities. This is the purpose of philosophy.

And Stoljar notes that those who do not accept the idea that staunch naturalism (our words) virtually requires acceptance of strict physicalism do not make a conceptual mistake—e.g., some minimalist version of libertarianism may well be conceptually consistent with faint-hearted naturalism (our words). Instead, says Stoljar, they are flying in the face of a thesis (warranted by science) we (allegedly) have overwhelming reason to believe.[8]

Consider also the titles of two books by the late Jaegwon Kim: *Mind in a Physical World* (1998) and *Physicalism, or Something Near Enough* (2005). In the former, Kim begins doing philosophy of mind with an explicit worldview commitment to staunch naturalism and strict physicalism. This commitment drives the book's argumentation, and the problem to be solved is attempting to discover a physicalist way of allowing for "mind," given staunch naturalist/ strict physicalism.

In the latter, the title invites the question: Near enough for what? Kim argues that we must embrace a dualist view on intrinsically characterized phenomenal qualia, specifically, sensations like pain. But this is close enough. What is it that Kim is trying to preserve, so his resulting view is close enough for that goal? Answer: It must be the kind of worldview or picture he feels comfortable living with, viz., physicalist naturalism, and his view is close enough to staunch naturalism and strict physicalism for his goal. Clearly, both books exhibit a prior commitment to staunch naturalism that informs his work in philosophy of mind. Elsewhere, Kim asserts as much, apparently with no sense that his claim will be controversial:

---

[7] Stoljar, D. (2010). *Physicalism*. London: Routledge, chapter one.

[8] For a powerful critique of the Standard Picture that actually reverses the epistemic and conceptual order of the Manifest and Scientific Image, see Robinson, H. (2016). *From the knowledge argument to mental substance* (pp. 133–144). Cambridge: Cambridge University Press.

> No serious competitor [to staunch physicalism] is visible on the horizon … Today, any proposed general ontology of the world, not just views about the mind-body relation, is defined by its relationship to materialism, the position that the world consists exclusively of bits of matter and structures made up of bits of matter, all behaving in accordance with physical law. Everything is an arrangement of matter, and living organisms and minded creatures are no exception.[9]

Because of the importance of a naturalistic worldview in the philosophy of mind and in the arguments in this book, we start our discussion with a presentation of important epistemic issues relevant to worldview evaluation, followed by a characterization of the inner logic of the most widely accepted version of naturalism (as noted above by Strawson), and we attempt to show that this version is what a naturalist *ought* to embrace—staunch naturalism. One can be a naturalist and a substance dualist (e.g., C. J. Ducasse) or a theist and a physicalist of one sort or another (e.g., Peter van Inwagen).[10] Accordingly, we offer arguments for Mere Substance Dualism that we take to stand on their own and be successful, irrespective of worldview considerations. If successful, most of these arguments present refuting or undercutting defeaters for monotheistic, especially Christian physicalists like Peter van Inwagen, Trenton Merricks, and Andrew Bailey.

Still, the worldview of naturalism—staunch or faint-hearted—funds various forms of physicalism for a majority of contemporary physicalists, and this factor is seldom analyzed in depth. We are not alone in this assessment. According to Andrea Lavazza and Howard Robinson:

> [T]he attacks (against dualism) are often dismissive, and the arguments are met without difficulty. Indeed, the present situation indicates that mental-physical dualism has fallen into disfavor in the contemporary intellectual scene not because the arguments for it have been found faulty but because it is assumed not to fit, or fit sufficiently well, with the naturalistic perspective that the advance of science has made seem so irresistible. One result of this has been a widespread failure to appreciate the force of those arguments.[11]

---

[9] Kim, J. (2018). Against Cartesian dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 154–167). Oxford: Wiley-Blackwell. The quote is on page 153.

[10] For a first-rate presentation and defense of an interesting version of physicalism in light of Abrahamic monotheism, see Baily, A. M. (2021). *Monotheism and human nature*. Cambridge: Cambridge University Press. For Christians, the nature of biblical anthropology will be important to one degree or another depending on one's understanding of biblical authority and intent. Still relevant is John Cooper's defense of ontological dualism and functional holism regarding human persons. See his *Body, soul and life everlasting* (2nd rev. ed.). (2000). Grand Rapids, MI: Eerdmans.

[11] Lavazza, A., & Robinson, H. (2016). Introduction. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (p. 5). London: Routledge. Parentheses ours.

Thus, it is important to keep an eye on the way (especially) naturalism factors into debates in philosophy of mind for this majority. As noted above, staunch naturalism entails strong physicalism and, among other things, eschews irreducible, ineliminable *sui generis* emergent properties. What we are calling faint-hearted naturalism embraces the existence of *sui generis* emergent properties. In light of this, we will offer a critique of emergent properties (systems, substances) after we discuss staunch naturalism.

Allegedly, sophisticated versions of panpsychism solve one of the problems with faint-hearted naturalism (see Section 2.4) that we present in our critique of emergent properties. For this and other reasons, the issue of panpsychism is important for assessing the success of an important part of our approach in the chapters to come. We will give panpsychism its proper due in Chapter 6.

2.2 IMPORTANT EPISTEMIC BACKGROUND ISSUES

While theism and naturalism are broad worldviews and not scientific theories, three issues that inform the adjudication between rival scientific theories are relevant to our topic.

2.2.1 Basicality

The first issue involves deciding whether it is appropriate to take some phenomenon as *basic* such that only a description and not an explanation for it is required, or whether that phenomenon should be understood as something to be explained in terms of *more basic* phenomena. For example, attempts to explain uniform inertial motion are disallowed in Newtonian mechanics because such motion is basic on this view, but an Aristotelian had to explain how or why a particular body exhibited uniform inertial motion. Thus, what is basic to one theory may be derivative in another.[12]

2.2.2 Naturalness

Issue two is the naturalness of a postulated entity in light of the overall theory of which it is a part. The types of entities postulated, along with the sorts of properties/powers they possess and the relations they enter should be “at home”—there should be a fittingness—with other entities in the theory. Here is our understanding of naturalness:

> *Naturalness*: Some entity (particular thing, process, categorical or dispositional property, or relation) *e* is natural for a theory T, just in case, either (i) *e* is a central, core entity of T or (ii) *e* bears a relevant similarity to central, core entities in *e*’s category within T.

For example, if *e* is in a category such as individual, force, property, event, relation, or cause, then *e* should bear a relevant similarity to central, core entities of T in that category.

---

[12] For our purposes here, we do not distinguish basicality and fundamentality. For more on this, see O’Connor, T. (2018). For emergent individualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 369–376). Oxford: Wiley-Blackwell.

This is a formal definition, and the material content given to it will depend on the theory in question.

Moreover, given rivals R and S, the postulation of *e* in R is *ad hoc* and *question-begging* against advocates of S if *e* bears a relevant similarity to the appropriate entities in S, and in this sense, is "at home" in S, but fails to bear this relevant similarity to the appropriate entities in R.[13] The notion of "being *ad hoc*" is notoriously difficult to specify precisely. It is usually characterized as an intellectually inappropriate adjustment of a theory whose sole epistemic justification is to save the theory from falsification. Such an adjustment involves adding a new supposition to a theory not already implied by its other features. In the context of evaluating rivals, R and S, the principle just mentioned provides a sufficient condition for the postulation of *e* to be *ad hoc* and question-begging.

The issue of *naturalness* is relevant to theory assessment between rivals in that it provides a criterion for advocates of a theory to claim that their rivals have begged the question against them or adjusted their theory in an inappropriate, *ad hoc* way. And though this need not be the case, naturalness can be related to *basicality* in this way:

> *Naturalness Basicality*: Naturalness can provide a means of deciding the relative merits of accepting theory R, which depicts phenomenon *e* as basic, vs. embracing S, which takes *e* to be explainable in more basic terms.

If *e* is natural in S but not in R, it will be difficult for advocates of R to justify the bald assertion that *e* is basic in R and that all proponents of R need to do is describe *e* and correlate it with other phenomena in R as opposed to explaining *e*. Such a claim by advocates of R will be even more problematic if S explains *e*.[14]

2.2.3 Explanatory Simplicity: Ockham's Razor and Ontological Parsimony

Two very different versions of simplicity are relevant to what follows. The first is *epistemic simplicity* (a.k.a. *Ockham's Razor*): Entities (assumptions, etc.) must not be multiplied beyond what is required to show that a theory (or proposition) has greater explanatory power or to confer more epistemic justification/warrant on that theory compared to rivals. This principle is far less controversial than the principle of ontological parsimony, and it follows from our ubiquitous understanding of explanation and justification/warrant. If a theory has more commitments than a rival, even though each does an equal job of explaining and receiving justification/warrant from the intellectual work they do, then the presence or absence of these extra commitments does no philosophical work and, therefore, is superfluous.

---

[13] For example, suppose theory S explains phenomena in terms of discrete corpuscles and actions by contact, while R uses continuous waves to explain phenomena. If some phenomenon *x* was best explained in corpuscularian categories, it would be *ad hoc* and question-begging for advocates of R simply to adjust their entities to take on particle properties in the case of *x*. Such properties would not bear a relevant similarity to other entities in R and would be more natural and at home in S.

[14] For example, suppose that R is Neo-Darwinism and S is a version of punctuated equilibrium theory. Simply for the sake of illustration, suppose further that R depicts evolutionary transitions from one species to another to involve running through a series of incrementally different transitional forms except for some specific transition *e* which is taken as a basic phenomenon, say, the discrete jump from amphibians to reptiles. S pictures evolutionary transitions in general, including *e*, as evolutionary jumps to be explained in certain ways that constitute S. In this case, given the presence of S, it would be hard for advocates of R to claim that their treatment of *e* is adequate against S. Phenomenon *e* clearly counts in favor of S over against R.

Still, there is a crucial point that must be kept in mind when employing epistemic simplicity: It is a tie-breaker with a *ceteris paribus* condition. Consequently, its application comes at the end of the evaluation of rivals. It cannot be employed at the beginning of such an investigation since, in this case, its application may rule out factors that, if allowed, would show that the rivals are not, in fact, equals, all things considered. In this way, epistemic simplicity distorts the evaluation of rivals. This is especially problematic when rivals satisfy different epistemic virtues.

The second principle of simplicity is *ontological parsimony*: We expect reality to be simple with respect to _____, where the blank is filled in differently by different philosophers (e.g., the form of our basic laws, the number of our basic laws, the number of categories and subcategories that carve up the world).

Some philosophers mistakenly conflate the two principles and move from one to the other as if they say the same thing. To illustrate, we select two notable and egregious examples of such conflation, beginning with Jaegwon Kim.[15] He begins his discussion of simplicity by focusing on its application to theories and hypotheses: “What can be done with fewer assumptions should not be done with more.” He goes on to say that we should “… posit no unnecessary entities—that is, entities that have no work to do.”[16] These statements express epistemic simplicity.

However, without warning, Kim shifts to an employment of ontological parsimony in arguing against embracing a huge number of irreducible psychoneural correlations. He asserts that the world depicted in such a view would contain a multitudinous, motley crowd of brute laws of nature, the very anthesis of simplicity. He concludes this argument by saying “We expect our basic laws to be reasonably simple, and reasonably few in number.”[17] Here Kim is making a claim that we expect reality itself to be simple. But ontological parsimony does not follow from epistemic simplicity.

A second illustration of this conflation comes from the pen of Andrew Bailey.[18] Bailey acknowledges that non-theists use what seems to be epistemic simplicity when they recommend in accordance with Ockham’s Razor that we slice off God as an unneeded ontological dangler, “… in favor of a less profligate (recklessly extravagant, wasteful, unneeded) naturalism.”[19] Without qualification, Ockham’s Razor is an epistemic principle. Bailey goes on to advance an argument based on “the conviction that simpler is better,” and the avoidance of “unattractive theoretical complexity.”[20] However, Bailey seems to shift without comment to ontological parsimony when he claims that monotheists have reasons to doubt that the world will or should be simple. Why? Because God is creative, and many theists claim that “God takes pleasure in the sheer variety of creation.”[21] Thus, the world will be a plenitude.

---

[15] See Kim, J. (2011). *The philosophy of mind* (3rd ed., pp. 98–102). Boulder, CO: Westview Press, especially 100.

[16] Ibid., 99.

[17] Ibid., 101.

[18] Bailey, *Monotheism and human nature*, especially 22–25.

[19] Ibid., 22. Cf. 27. Parentheses ours.

[20] Ibid., 22. Cf. the bottom of 23.

[21] Ibid., 24.

While we accept epistemic simplicity, we agree with many philosophers who hold that ontological parsimony is controversial. Indeed, we do not accept it. We see no reason to think that before we engage in metaphysics, we must be saddled with *ontological parsimony*. We reject a burden of proof placed on rivals by those employing *ontological parsimony* in our deliberations about accepting, e.g., abstract objects, mental entities, the object reality of irreducible secondary qualities. This is so because we adopt a shopping-list metaphysical methodology and examine the relevant arguments on their own merits in these and other areas with no need to find the world itself to be simple.

However, as we will now discover, we do think that a contemporary scientific/philosophical naturalist should be a staunch naturalist and strong physicalist. And the best metaphysical approach for such thinkers should be the employment of serious metaphysics, especially since the naturalist must solve the location problem within the constraints of the naturalist epistemological commitments and the Grand Story. To these issues we now turn.

2.3 CONTEMPORARY SCIENTIFIC NATURALISTS SHOULD BE STAUNCH NATURALISTS AND STRONG PHYSICALISTS

Clearly, there are many versions of naturalism.[22] However, if we look carefully at the overall structure of contemporary scientific naturalism as a worldview, we will gain insight into why staunch naturalism is—and should be—the ubiquitously preferred view.[23] Contemporary scientific naturalism usually includes:

1. different aspects of a naturalist epistemic attitude (for example, a rejection of so-called "first philosophy" along with an acceptance of either weak or strong scientism);
2. a Grand Story, which amounts to an etiological account of how all entities whatsoever have come to be, told in terms of an event-causal story described in natural scientific terms with a central role given to the atomic theory of matter and evolutionary biology;
3. a general ontology in which the only entities allowed are those that either (a) bear a relevant similarity to those thought to characterize a completed form of physics (staunch naturalism or strong physicalism) or (b) are wholly dependent on and determined by the entities of microphysics and can be explained according to a causal necessitation (nomological or metaphysical) requirement (given a "suitable" arrangement of matter, the emergent entity *must* arise) in terms of the Grand Story and the naturalist epistemic attitude (weak naturalism or physicalism[24]).

---

[22] See Levy, N. (2016). Naturalism and free will. In K. J. Clark (Ed.), *The Blackwell companion to naturalism* (pp. 305–318). Malden, MA: John Wiley and Sons, Inc.; cf. Horst, S. (2009). Naturalisms in philosophy of mind. *Philosophy Compass*, *4*/1, 219–254.

[23] For a good example of a naturalist who understands this and, consequently, advocates for staunch naturalism, see Rosenberg, A. (2015). Disenchanted naturalism. *Kritikos*, *12* (January–April). https://intertheory.org/rosenberg.htm.

[24] The majority of naturalists adopt a metaphysical or nomological necessitation requirement for the emergent entity relative to the base from which it originates and is sustained. However, some demure. For a good example of this debate, see Searle, J. (1992). *The rediscovery of the mind* (pp. 100–105). Cambridge, MA: MIT Press.

2.3.1 The Naturalist Epistemic Attitude

2.3.1.1 Epistemology and the Origin and Nature of Human Noetic Faculties

For most naturalists, the ordering of these three ingredients is important. The naturalist epistemic attitude serves as justification for the naturalist etiology, which, in turn, helps to justify the naturalist's ontological commitment. Moreover, naturalism seems to require coherence among the postulates of these three different facets of the naturalistic turn. David Papineau claims that we should set philosophy within science in that philosophical investigation should be conducted within the framework of our best empirical theories. It follows that "… the task of the philosophers is to bring coherence and order to the set of assumptions we use to explain the empirical world."[25] For example, there should be coherence among third-person scientific ways of knowing; a physical, evolutionary account of how our sensory and cognitive processes came to be; and an ontological analysis of those processes themselves. As Carlo Penco, Michael Beanley, and Massimiliano Vignolo put it, "The root of naturalism is the empiricist view that our cognitive capacities must be explained in accord with the idea of ourselves as physical organisms that bear causal relations to the environment."[26]

Any entities that are taken to exist should bear a relevant similarity to entities that characterize our best physical theories; their coming-to-be should be intelligible in light of the naturalist causal story; and they should be knowable by scientific means. Put differently, any candidate entity must be locatable in the Grand Story and, in turn, known by way of the naturalist epistemic attitude.

This ordering of these non-capricious, widely accepted features of naturalism put naturalists on something akin to the horns of a dilemma: (1) Embrace a staunch ontology with the hope of retaining superior epistemic warrant/explanatory power by adhering to the ordering listed above, and embark on the difficult and highly counterintuitive task of reducing or eliminating entities that are recalcitrant in order to locate them in the Grand Story; or (2) Embrace a weak naturalist ontology, leave certain commonsense entities as they are, but lose the claim of superior epistemic warrant/explanatory power by (i) violating epistemic simplicity and ontological parsimony; (ii) bloating the naturalist ontology such that numerous entities must be taken as utterly inexplicable brute facts since they cannot be located in accordance with the Grand Story and accepted epistemology; (iii) accepting unnatural entities via inappropriate *ad hoc* adjustments to a naturalist ontology that beg the question against a theistic worldview in which such entities—especially those related to philosophical anthropology—are at home and easily located.[27] Naturalists who deviate from this epistemic commitment owe us an account of how other ways of knowing (*a priori* awareness, the first-person point of view) could arise from purely physical processes constitutive of evolution, and be located in the Grand Story.

---

[25] Papineau, D. (1993). *Philosophical naturalism* (p. 3). Oxford: Blackwell.

[26] Penco, C., Beanley, M., & Vignolo, M. (2007). Introduction. In C. Penco, M. Beanley, & M. Vignolo (Eds.), *Explaining the mental: Naturalist and non-naturalist approaches to mental acts and processes* (pp. xxiv–xxv). Newcastle, UK: Cambridge Scholars Publishing.

[27] For a recent argument showing that attempts to embrace naturalism without strong physicalism are failures, see Robinson, H. (2016). *From the knowledge argument to mental substance* (pp. 104–111). Cambridge: Cambridge University Press.

2.3.1.2 Scientism

Despite protests to the contrary, scientism constitutes the core of the naturalist epistemology.[28] Long ago, Wilfrid Sellars said that "in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not."[29] Most contemporary naturalists embrace either weak or strong scientism. According to the former, nonscientific fields are not worthless, nor do they offer no intellectual results, but they are vastly inferior to science in their epistemic standing and do not merit full credence. According to the latter, unqualified cognitive value resides in science and in nothing else. Either way, naturalists are incredibly skeptical of any claims about reality that are not justified by scientific methodology in the hard sciences.

2.3.1.3 Third-Person Ways of Knowing

Moreover, that methodology is a third-person one that sanctions only entities capable of exhaustive description from a third-person perspective. Skepticism prevails for entities that require the first-person perspective as their basic mode of epistemic access.[30] For such naturalists, the exhaustive or elevated nature of scientific knowledge entails that either the only explanations that count or the ones with superior, unqualified acceptance are those employed in the hard sciences.[31]

At least two philosophical theses elaborate the naturalistic epistemic and methodological constraints for philosophy. First, there is no such thing as first philosophy (a subject matter in philosophy that is autonomous from and authoritative in comparison to science); rather, there is continuity between philosophy and natural science. As naturalist Patricia Churchland puts it, "Naturalism follows hard upon the heels of the understanding that there is no first philosophy."[32] Second, scientific theories that are paradigm cases of epistemic/explanatory success, e.g., the atomic theory of matter, evolutionary biology, employ combinatorial modes of explanation. Thus, any process that constitutes the Grand Story and the coming-to-be and nature of any entity in the naturalist ontology should exhibit an ontological structure analyzable in terms that are isomorphic with such modes of explanation.

2.3.1.4 Combinatorial Modes of Explanation

Colin McGinn has defended this idea along with what he takes it to entail, viz., the inability of naturalism to explain genuinely unique emergent properties:

> Can we gain any deeper insight into what makes the problem of consciousness run against the grain of our thinking? Are our modes of theorizing about the world of the wrong shape to extend to the nature of mind? I think we can discern a characteristic structure possessed by successful scientific theories, a structure

---

[28] See Moreland, J. P. (2018). *Scientism and secularism: Learning to respond to a dangerous ideology.* Wheaton, IL: Crossway.

[29] Sellars, W. (1963). *Science, perception, and reality* (p. 173). London: Routledge and Kegan Paul.

[30] Churchland, P. (2013). *Matter and consciousness* (pp. 24–25, 120–131).

[31] We are assuming here a realist construal of explanation.

[32] Churchland, P. (1986). *Neurophilosophy: Towards a unified science of the mind/brain* (p. 277). Cambridge, MA: MIT Press.

> that is unsuitable for explaining consciousness …. Perhaps the most basic aspect of thought is the operation of *combination*. This is the way in which we think of complex entities as resulting from the arrangement of simpler parts. There are three aspects to this basic idea: the atoms we start with, the laws we use to combine them, and the resulting complexes … I think it is clear that this mode of understanding is central … [and] our scientific faculty involves representing the world in this combinatorial style.[33]

Accordingly, J. J. C. Smart paradigmatically noted: "It seems to me that science is increasingly giving us a viewpoint whereby organisms are able to be seen as physicochemical mechanisms … There does seem to be, so far as science is concerned, nothing in the world but increasingly complex arrangements of physical constituents. All except for one place: in consciousness … I just cannot believe that this can be so. That everything should be explicable in terms of physics … except the occurrence of sensations seems to me to be frankly "unbelieveable."[34]

Along similar lines, Paul Churchland claims that the existence of emergent phenomenal properties

> … sits poorly with the joint claim that mental properties emerge from nothing more than the organizational achievements of physical matter. If that is how mental properties are produced, then one would expect a physical account of them to be possible, even inevitable. The simultaneous claim of evolutionary emergence and irreducibility is prima facie puzzling, if not self-contradictory.[35]

In sum, the naturalist epistemic attitude countenances and countenances only (1) the hard sciences along with mathematical empiricism as the paradigm of knowledge, (2) a certification of third-person ways of knowing while eschewing the first-person, (3) an employment of combinatorial modes of explanation for properties which will all be structural and for the nature, coming-to-be, and perishing of all macro-wholes above the basic level of micro-physical particles (waves, wavicles, strings, turtles or whatever; hereafter, simply "particles") and (4) a rejection of first philosophy. So, for example, given that an exercise of active power is a non-relational, intrinsically characterizable mental event, and the property of being active power is a simple quality and not a structural property composed of separable parts, it would seem that these entities at the heart of libertarian free will are ruled out by the naturalist epistemic attitude and the Grand Story it certifies. We examine this topic in detail in Chapter 9.

---

[33] McGinn, C. (1999). *The mysterious flame* (pp. 55–56). New York, NY: Basic Books. Cf. 54–62, 90, 95. The same point is made by Richard Swinburne. See his *The evolution of the soul* (pp. 3–4). (1986). Oxford: Clarendon Press.

[34] Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, *68*, 141–156.

[35] Churchland, P. *Matter and consciousness* (pp. 19–20). Though he is a faint-hearted naturalist with respect to emergent moral properties, Russ Shafer-Landau acknowledges that staunch naturalism (he calls it "classic naturalism") had been the default view because naturalists take this position to be forced on them by the ontological constraints of science coupled with epistemic simplicity. Shafer-Landau admits that he lacks a knock-down argument against staunch naturalism. See Shafer-Landau, R. (2003). *Moral realism: A defence* (pp. 65–72). Oxford: Clarendon Press.

2.3.2 The Grand Story

2.3.2.1 Précis of the Grand Story

The naturalist has an account of how all things whatever came to be. Contemporary scientific naturalists boast that this account and only this account is certified by naturalist epistemic commitments. Let us call this account *the Grand Story*. The details of the Grand Story need not concern us here. Suffice it to say that some version of the Big Bang is the most reasonable view currently available. On this view, all of reality—space, time, and matter—came from the original “creation” event and the various galaxies, stars, and other heavenly bodies eventually developed as the expanding universe went through various stages. On at least one of those heavenly bodies—earth—some sort of pre-biotic soup scenario explains how living things came into being from nonliving chemicals. And the processes of evolution, understood in either neo-Darwinian or punctuated equilibrium terms, gave rise to all the life forms we see, including human beings. Thus, all organisms and their parts exist and are what they are because they contributed to (or at least did not hinder) the struggle for reproductive advantage; more specifically, because they contributed to the tasks of feeding, fighting, fleeing, and reproducing.

2.3.2.2 Four Core Elements to the Grand Story

There are four important things to note about the Grand Story. First, in keeping with the naturalist epistemic attitude, at the core of the Grand Story are two theories that countenance only combinatorial modes of explanation: the atomic theory of matter and evolutionary theory. Any appeal to the existence of emergent properties violates this ontological constraint.

Second, it is an expression of a scientistic version of philosophical monism according to which everything that exists or happens in the world is (in principle) susceptible to explanations by natural scientific methods in the hard sciences. Whatever exists or happens in the world is natural in this sense. Prima facie, the most consistent way to understand naturalism in this regard is to see it as entailing some version of strong physicalism: everything that exists is fundamentally matter, most likely, elementary “particles” (whether taken as points of potentiality, centers of mass/energy, units of spatially extended stuff/waves or reduced to/eliminated in favor of fields), organized in various ways according to the laws of nature.[36] By keeping track of these particles and their physical traits we are keeping track of everything that exists. No nonphysical entities exist, including emergent ones.

This constitutes a strong sense of physicalism and staunch naturalism. When naturalists venture away from strong physicalism, however, they still argue that additions to a strong physicalist ontology must be depicted as wholly rooted in, emergent from, dependent upon the physical states and events of the Grand Story. These states/events are in principle wholly describable in the language of physics. This is weak physicalism and faint-hearted naturalism. Later, we will say more about why emergent properties should not be allowed in a naturalist ontology.

---

[36] We will continue to talk in terms of particles and not fields because much of the literature in philosophy of mind does so, e.g., debates about atomic simples and constitution, but we do not think anything important hangs on this. Cf. Clifton, R., & Halverson, H. (2002). No place for particles in relativistic quantum theories? *Philosophy of Science*, *69*, 1–28.

Third, because of the importance of laws in science and of combinatorial processes in causation, the Grand Story is constituted by event causality and eschews both irreducible teleology and agent causation in which the first relatum of the causal relation is in the category of substance and not event. And the Grand Story is deterministic in two senses: diachronically such that the state of the universe at any time *t* coupled with the laws of nature determine or fix the chances for the state of the universe at subsequent times; synchronically such that the features of and changes regarding macro-wholes are wholly dependent on and determined by micro-physical phenomena.

2.3.3 The Standard Naturalist Ontology

2.3.3.1 The Location Problem

For present purposes, it is important to say a bit more about the criteria for naturalist ontological commitments. A good place to start is with what Frank Jackson calls the location problem.[37] According to Jackson, on the basis of the superiority of scientific ways of knowing exemplified by the hard sciences, naturalists are committed to a fairly widely-accepted physical story about how things came to be (the Grand Story) and what they are.[38] Given these commitments, the location problem is the task of locating or finding a place for some entity (for example, semantic contents, mind, agency) in that story.

For Jackson, the naturalist must either locate a problematic entity in the basic story or eliminate the entity. Roughly, an entity is located in the basic story just in case it is entailed by that story. Otherwise, the entity must be eliminated. Jackson correctly notes that a naturalist ought to adopt serious metaphysics: His or her ontology should start with the smallest number of different sorts of entities—those we find in our best theories in physics—and incorporate those entities that can be explained by the Grand Story's combinatorial processes, viz., mereological aggregates for individuals, structural properties for attributes, external relations whose relata are separable parts.

It is hard to see how the ontology generated by the combinatorial processes at the heart of scientific explanation, e.g., as seen in the atomic theory of matter and evolutionary biology, could countenance simple emergent properties or substances. As we will explain below, we think that "emergence" is just a name for the problem to be solved (How could simple emergent properties and substances emerge if you start with particles as depicted by physics and just rearrange them over time?), and not a solution.

Among other things, this means that without some pretty serious, wildly *ad hoc* adjustments, the sort of unity possessed by consciousness (and, perhaps, its ground) cannot be located or otherwise explained, given staunch naturalism and strong physicalism. Jackson grasps the connection between accepting the epistemic superiority of naturalism and its Grand Story and deciding between weak and strong physicalism. For Jackson, *if naturalism is to have superior explanatory power, this entails strong physicalism/staunch naturalism.*

---

[37] Jackson, F. (1998). *From metaphysics to ethics* (pp. 1–5). Oxford: Clarendon Press.

[38] For an excellent précis of the Grand Story, see McGinn, *The mysterious flame* (pp. 12–18).

2.3.3.2 Microphysical Priority and the Mereological Hierarchy

Perhaps the most important aspect of the staunch naturalist ontology is the mereological hierarchy. Let us construe this hierarchy in terms of individual entities and properties rather than in terms of concepts or linguistic descriptions. So understood, the standard mereological hierarchy consists of an ascending level of entities in the category of individual such that for each level above the ground level of elementary micro-physics (at which entities have no further physically significant separable parts), wholes at that level are composed of the separable parts at lower levels standing in external relations to each other.

Thus, from bottom to top, we get micro-physical entities (strings, waves, particles, fields?), subatomic particles, atoms, molecules, cells, living organisms, ecosystems, and so on. The relationship between individuals at level n and n+1 is the separable part/whole relation. In the category of property, structurally supervenient properties—those that are merely a new relational combination of the parts and properties at the subvenient level—such as being H_2O—fit naturally into the naturalist combinatorial depiction of how all things have developed since the Big Bang.

It is very important to realize that the naturalist ontology and the mereological hierarchy just described are not arbitrary commitments that come out of thin air. They follow from the very inner logic of naturalism that (allegedly) gives it epistemic/explanatory superiority over rival worldviews. Arguably, the late Jaegwon Kim was among the top philosophical naturalists of the last fifty years. Read carefully this claim he made without thinking it needed defending:

Today, any proposed general ontology of the world … is defined by its relation to materialism, the position that the world consists exclusively of bits of matter and structures made up of bits of matter, all behaving in accordance with physical law. Everything is an arrangement of matter and living organisms and minded creatures are no exceptions.[39]

In our view, it borders on an intellectual dereliction of duty to embrace faint-hearted naturalism and quantify over emergent entities. But perhaps this judgment is premature. How does naturalistic emergence fare in light of Kim's remark and objections raised against it?

2.4 NATURALISTIC EMERGENCE

For those who accept the existence of genuine, irreducible, nonphysical conscious states and, at the same time, affirm naturalism, the most popular current strategy is to adopt faint-hearted naturalism and weak (usually, nonreductive) physicalism, and appeal to some sort on natural emergence. According to the staunch-naturalist account, matter at the most basic level—the microphysical level—is or will be completely describable within an ideal physics. Staunch naturalist David Papineau's view on this matter is canonical. While acknowledging that progress toward an ideal science may require categorial supplementation of some sort, he maintains the following: "What I do not believe is that they [the current categories of physics plus future supplementation] will need supplementation by psychological categories."[40] [brackets ours]

---

[39] Kim, J. (2018). Against Cartesian dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 152–167). Oxford: Wiley-Blackwell. The quote is on page 153.

[40] Papineau, 31.

That ideal description will be able to capture the complete nature of basic matter, and it will entail that basic matter is bereft of any mental entities, including mental potentialities. So far, a faint-hearted naturalist would agree. But they go on to affirm that when matter reaches a certain, “appropriate” level of complexity, a completely new range of *sui generis* emergent properties, especially conscious properties, appears from a subvenient base wholly bereft of, say, consciousness or conscious potentialities.

In the contemporary literature, this emergence strategy takes two forms: a strictly naturalist account and some sort of panpsychist view. We will discuss the naturalist account in this section and provide a treatment of panpsychism in Chapter 6.

However, it will be helpful to begin our discussion with some important distinctions. This will allow us to clarify in more detail precisely what an emergent property is.

2.4.1 Three Important Distinctions

2.4.1.1 Weak Vs. Strong Property Supervenience

Three distinctions are relevant for our purposes. The first is between weak (WPS) and strong property supervenience (SPS). *Weak supervenience* compares individuals in the *same* possible world: if in the actual world, two cases agree in all subvening respects, the two will also agree in all supervening respects. *Strong supervenience* compares individuals within or *across* possible worlds: given two individuals in different counterfactual contexts, but themselves agreeing in all subvenient respects, the two will also agree in all supervening respects.[41]

Two comments: we examine only property supervenience; the modality of strong supervenience is nomological or metaphysical.

2.4.1.2 Emergent Vs. Structural Supervenience

The second distinction is between emergent and structural properties/supervenience. *Emergent properties* (EPs) are completely unique, new kinds of properties different from those that characterize their subvenient bases. Accordingly, and minimally, *emergent supervenience* (ES) (or simply, emergence) depicts the supervenient property as a simple, novel property different from and not composed of subvenient entities. Being painful is an example.

Keep in mind that in the first half of the 20th century, emergent properties were defined epistemologically:[42]

> *Emergent Property*: Property P is an emergent property of some particular *x* at level l_n just in case (i) P is a property of *x* at l_n, and (ii) no amount of knowledge of (or descriptive statements about) entities at subvenient levels below l_n would justify a prediction of (or logically entails a descriptive statement about) P at l_n.

---

[41] Tiller, P. (1995). Supervenience. In J. Kim & E. Sosa (Eds.), *The Blackwell companion to metaphysics* (p. 485). Oxford: Basil Blackwell. For a presentation of a number of important distinctions essential to understanding supervenience, see Moreland, J. P. (2020). Wielenberg and emergence: Borrowed capital on the cheap. In A. L. Johnson (Ed.), *God and morality: What is the best account of objective moral values and duties?* (pp. 93–114). New York, NY: Routledge.

[42] For a classic comparison of the epistemic and ontological characterizations of emergent properties, see Nagel, E. (1979). *The structure of science* (pp. 366–380). Indianapolis: Hackett.

In this sense an emergent property (or a statement about it) is surprising, unexpected, and utterly brute (we use “brute” or “brute fact” to refer to a brute thing/entity and not a brute truth) relative to knowledge of (or statements about) lower levels. In principle, a brute fact has no explanation for why it obtained. Thus, the prior probability of emergent properties, given complete knowledge of lower levels, would be zero.

By contrast, a structural property is constituted by a configurational pattern of subvenient features. More specifically,

> *Structural Property* (SP): P is a structural property of some particular *x* just in case P is constituted by a configurational pattern of P’s subvenient parts, properties, relations, and events.

A structural property is not a new *kind* of property; it is merely a new pattern of subvenient entities. Some characterize emergent and structural supervenience (SS) as causal and constitutive, respectively. We will use the notion of an emergent property as merely being a mereologically simple, novel, *sui generis* property. From this depiction alone, an emergent property could be epiphenomenal or have so-called top-down causal influence.

2.4.1.3 Making Precise the Notion of an Emergent Property

Third, what many call *weak emergence* (WE) and *strong emergence* (SE) maps fairly well onto what we are calling structural supervenience and emergent supervenience, respectively.[43] WE entails ontological and causal reduction, while SE entails ontological novelty. There is some debate about whether a strongly emergent property must have new causal powers (or be involved with new laws), but most think it does.[44] Moreover, a crucial aspect of (strongly) emergent properties is that they arise out of, are continually sustained by and wholly dependent on, and do not change the nature of their unperturbed base entities. Thus, *SE properties (hereafter, emergent properties) are novel, mereologically simple, nonstructural, fundamental properties that exhibit new causal powers or conform to new causal laws and wholly arise out of and are sustained by base entities (and their relational structure) unaffected by and devoid even of the potentiality for their emergence.*

2.4.2 Additional Metaphysical Aspects of Emergent Properties

2.4.2.1 Twin Hallmarks of Emergence

To begin with, Mark Bedau identifies Twin Hallmarks of emergence. All emergence involves macrophenomena that (i) arise from and depend on more basic phenomena, and (ii) are simultaneously autonomous from the microlevel on which they depend. Consequently, a conception of emergence requires an interpretation of dependence and autonomy that is metaphysically coherent, without negative metaphysical implications, and that helps scientists understand some interesting aspect of reality. Bedau asserts that all the evidence shows that SE is scientifically irrelevant, and it straps us with a host of brute, inexplicable phenomena that should be avoided if at all possible. Moreover, it seems that all alleged cases of SE can be explained away.[45]

---

[43] Bedau, M. A. (2010). Weak emergence and context-sensitive reduction. In A. Corradini & T. O’Connor (Eds.), *Emergence in science and philosophy* (pp. 46–63). New York: Routledge.

[44] See O’Connor, T. (2000). *Persons and causes* (pp. 110–115). New York, NY: Oxford University Press.

2.4.2.2 Seven Characteristics of Emergent Properties

Mariusz Tabaczek adds further insights about the metaphysical characteristics of emergent properties:[46]

1. Nonadditivity of the sum of causes of a whole's parts.
2. Novelty of complex processes, entities, and properties.
3. Emergence included ontological dependency on base entities that goes beyond mere constraints of property covariance proffered by property supervenience.
4. The mereological hierarchy as a backdrop with higher levels of complexity involving holistic fundamental systems with emergent properties *sans* part-priority.
5. The presence of emergent laws.
6. Nondeducibility, nonpredictability, and irreducibility of emergent.
7. Emergent properties exhibit "downward" causation.

Given these clarifications, we turn to an exploration of the problems that emergent properties present for advocates of faint-hearted naturalism.

2.4.3 Confusion About the Basic Problem of Emergent Properties and Faint-Hearted Naturalism

The mereological hierarchy coupled with part-priority—usually microphysical part-priority—will be examined in detail in Chapter 7. But for now, it is important to identify exactly why these commitments are so important to scientific naturalism. Many argue that these commitments derive from (1) the fact that physicalism, including the causal closure of the micro-physical, requires them, as do (2) the unrivaled success of microphysical explanation, and (3) mereological Mooreanism (it is a Moorean fact that physical composite objects are composed by and only by separable micro-physical parts).[47]

However true these claims may be, they are subject to the criticism that organisms and other higher-level dynamic systems seem to undermine micro-physical part-priority or reduction, and emergent properties, processes, and wholes are needed to understand these entities properly. As interesting as this debate is (and we will enter it in Chapter 7), we think that the reasons offered for microphysical part-priority (and against emergent properties) are not the primary grounds for eschewing emergent properties, given scientific naturalism. We think those reasons are threefold, especially regarding emergent mental properties and substances.

First, the naturalist epistemic commitments (e.g., commitment to third-person ways of knowing and eschewing first-person knowledge), along with the combinatorial processes of the Grand Story, make the "location" of emergent (especially conscious) entities almost impossible. In this regard, Paul Churchland correctly notes:

---

[45] Bedau, 47–51.

[46] Tabaczek, M. (2019). *Emergence* (pp. 45–61). Notre Dame, IN: Notre Dame Press.

[47] See Inman, R. (2019). *Substance and the fundamentality of the familiar* (pp. 122–130). New York, NY: Routledge.

> [T]he important point about the standard evolutionary story is that the human species and all of its features are the wholly physical outcome of a wholly physical process … If this is the correct account of our origins, then there is neither need, nor room, to fit any nonphysical substances or properties into our scientific accounts of ourselves. We are creatures of matter. And we should learn to live with that fact.[48]

Second, given a rival theistic worldview and the adjudicative criteria of basicality and naturalness, it is question-begging and inappropriately *ad hoc* simply to announce as a brute fact (entity) the existence of (especially conscious) emergent entities.[49]

Finally, emergent entities violate the principle of *ontological parsimony*, thereby weakening a naturalist worldview from claiming to have superior explanatory success *vis a vis* its rivals. As Jaegwon Kim put it:

> [A]ll such correlations would have to be taken as "brute" basic laws of the world—"brute" in the sense that they are not further explainable and must be taken to be among the fundamental laws of our total theory of the world … But such a theory of the world should strike us as intolerably complex and bloated—the very antithesis of simplicity and elegance we strive for in science.[50]

Besides being a *per se* violation of *ontological parsimony*, the emergence-correlation move leaves a multitude of examples—"a huge and motley crowd" as Kim put it—of unanswered questions that weakens the alleged epistemic/explanatory superiority claimed by naturalists (e.g., Why emergent property P and not Q? Why any emergent property at all?). David Papineau astutely recognized this problem. In speaking of various emergentist physicalist "theories of consciousness" that attempt to explain why irreducible phenomenal conscious states regularly correlated with some physical states and not others, Papineau rightly observes that for the naturalist,

> … a natural reaction to all such theories is that they simply fail to address the philosophical question at issue. At best such a theory will specify some structural or other physically acceptable characteristic (A, say) which is co-extensive with the class of states we are pre-theoretically inclined to count as conscious. But then we still seem to face the question: *why* does consciousness emerge in just those cases? And to this question, physicalist "theories of consciousness" seem to provide no answer.[51]

---

[48] Churchland, *Matter and consciousness* (p. 35).

[49] For more on this, see Moreland, J. P. (2011). Oppy on the argument from consciousness: A rejoinder. *European Journal for Philosophy of Religion*, *3*(1), 213–226.

[50] Kim, J. (2011). *Philosophy of mind* (3rd ed., pp. 100–101). Boulder, CO: Westview Press.

[51] Papineau, D. (1993). *Philosophical naturalism* (p. 119). Oxford: Blackwell. See also Papineau, D. (2006). Comments on Galen Strawson. In A. Freeman (Ed.), *Consciousness and its place in nature* (pp. 100–109). Exeter, UK: Imprint Academic; *Thinking about consciousness* (pp. 253–55). (2002). Oxford: Clarendon Press.

Papineau goes on the say that physicalists should simply reject this question by identifying A with being conscious, asserting that there are not two properties here. Being conscious just is having A. In our terms, Papineau's recommendation is that naturalists stick with the staunch version.

2.4.4 Classification of Emergent Properties and a New Difficulty for Faint-Hearted Naturalism

As already noted, emergent properties simply do not belong in a naturalist ontology, as observed above by Mark Bedau. Besides this, there is a further problem for faint-hearted naturalists presented by emergent properties that becomes clear when we categorize different kinds of (so-called) emergent properties. To see this, consider the following:

Emergence$_0$: New features that can be deduced from base (e.g., fractal patterns).
Emergence$_1$: Ordinary structural properties (e.g., being water, solidity).
Emergence$_{2a}$: *Sui generis*, simple, intrinsically characterizable, new kinds of properties relative to base that are also epiphenomenal (e.g., being painful construed epiphenomenally).
Emergence$_{2b}$: *Sui generis*, simple, intrinsically characterizable, new kinds of properties relative to base with new causal powers construed as passive liabilities (e.g., being painful understood as having top/down causal liabilities).
Emergence$_{2c}$: *Sui generis*, simple, intrinsically characterizable, new kinds of properties with active power (e.g., being active power that characterizes most versions of agent causation).
Emergence$_3$: An emergent, suitably unified mental subject or I with active power.

Clearly, emergence$_0$ and emergence$_1$ fit nicely in the mereological hierarchy and conform to the naturalist epistemology (e.g., combinatorial explanation) and Grand Story. These are actually structural and not emergent properties. But emergence$_{2a}$ through emergence$_3$ should be disallowed for various reasons, e.g., they are hard to locate within the epistemic and Grand-Story aspects of naturalism, they resist functionalization. It would seem that all a naturalist could do with them is simply to label them as contingent brute facts and assert that they are not a problem for the naturalist. Apart from the results of libertarian free actions, contingent brute facts are problematic precisely due to their contingency. One need not embrace strong versions of the Principle of Sufficient Reason to see that since a contingent brute fact could have failed to obtain, one is left with an obviously pressing question as to why it did. This is not true for necessary beings that are brute facts, but contingent brute facts cry out for explanation.

Moreover, there is an increasingly heavy burden of proof on a naturalist ontology as one moves from emergence$_{2a}$ to emergence$_3$. All types of emergence fall prey to previous arguments against emergent entities as we will see. Emergence$_{2a}$ requires less justification than stronger forms of emergence because it does not require a rejection of closure. Emergent$_{2b}$ is subject to these arguments and additional difficulties with top/down causation and causal closure. But relative to emergence$_{2c}$ and emergence$_{3,}$ it has the advantage of exhibiting the same sort of causal power—passive liability subject to law—that characterizes causal particulars at the microphysical level.

Emergence$_3$ has all the problems exemplified by emergent$_{2b,}$ and it also suffers from having a completely unique sort of causal power—active power—different from the causal powers that range throughout the naturalist ontology outside of agent-causal events. Emergence$_3$ also suffers from a further fact not easily accommodated in the naturalist ontology: the indexical fact associated with “I” and difficulties with explaining how one can get a sort of primitive, substantial unity in which its various inseparable parts/faculties are internally related to the substantial subject from a mereological aggregate constituted by a structural arrangement of separable parts that stand in external relations to each other and their mereological whole.

2.4.5 Eight Additional Criticisms

2.4.5.1 *“Emergence” Is Just a Label*

For one thing, an appeal to emergent properties has always seemed suspect to us if it is offered as an explanation: *“emergence” is not a solution but a name for the problem to be solved*. It is a label and that’s all. This is especially problematic in the current setting in which scientists and philosophers of science are increasingly elevating the importance of seeking mechanisms that explain how and why emergent properties appear under a specific set of base conditions.[52]

In a closely analogous context, Jaegwon Kim has convincingly made the same charge against the widely employed, but illicit inclusion of dependency in strong property supervenience.[53] He notes that such supervenience is merely a principle of property covariance and, as such, it is consistent with a number of disparate views on the mind/body problem, including some forms of property or substance dualism. According to Kim, supervenience is a name that provides no explanation of the problem to be solved, viz., the dependency and determination of the mental on or by the physical.

Kim’s solution to this problem is to adopt a version of physicalist functional realizationism, which entails that mental properties supervene on physical properties because physical properties realize mental properties. The point here is not to evaluate Kim’s solution; instead, it is to acknowledge his charge that slapping the label “supervenience” on the problem of mental dependency on the physical is no solution at all. Thus, Kim recognized that further theoretical work is needed to satisfy that need.

We make the same claim about the naturalist employment of “emergence.” We recognize that supervenience *tout court* is not precisely analogous to genuine emergence. Still, we think that Kim’s treatment of strong property supervenience *is* appropriately analogous to our claim about “emergence.” Employment of “emergence” is more egregious than using “supervenience” in the Kim discussion. Why? As Kim shows, at least the “supervenience” case allows room for developing an explanatory theory that grounds and explains why supervenient dependency is true.

But this is not possible for emergent properties. Recall that until the middle of the 20th century, emergent properties were characterized epistemically, e.g., as being unpredictable from an exhaustive knowledge of base entities. Genuine emergence is the Platonic Form of being a brute fact; as such, there is no deeper explanation of emergent properties on offer. Consequently, the charge that emergence is just a name for the problem and not a solution seems justified.

---

[52] See Tabaczek, pp. 27–34.

[53] Kim, J. (1998). *Mind in a physical world* (pp. 1–27). Cambridge, MA: MIT Press. Kim employs nomological necessity in his understanding of the sort of modality implied by “strong.”

In the present context, while we disagree with his overall position, it is instructive to notice an important aspect of Colin McGinn's Mysterian Naturalism.[54] According to McGinn, given that there is no knowable, current naturalistic solution to the origin of consciousness or its regular correlation with matter, and due to our cognitive limitations, he sets out two conditions for any solution to these issues: (i) It must be a naturalistic solution. (ii) It must depict the emergence of consciousness and its regular correlation with matter as *necessary and not contingent* facts.

More specifically, there must be three kinds of unknowable natural properties that solve the problem. First, there must be some general properties of matter that enter into the production of consciousness when assembled into a brain. Thus, all matter has the potentiality to underlie consciousness. Second, there must be some natural property of the brain he calls C* that unleashes these general properties under the right conditions. Third, just as the brain must have a hidden unknowable structure that allows consciousness to emerge from it, so consciousness must have a hidden unknowable essence that allows it to be embedded in the brain.

The important point about McGinn's project is this: He recognized that a naturalist employment of "emergence" to solve the problems of the origin of consciousness and its regular correlation with matter is no explanation at all. It is what needs to be explained, and on this basis, McGinn offers a solution for the label. The simple fact that he does this demonstrated his recognition of the explanatory vacuity of the label "emergence."

2.4.5.2 Emergence and Empirically Equivalent Theories

Second, as a label, it leaves open other empirically equivalent views, viz., that the new kind of property may reside in a new substance (see Section 2.4.5.8) or be due to God's regular intention that the property appear under the same set of circumstances. As Timothy O'Connor has noted, if the lawlike link between occurrent subvenient properties and their associated occurrent emergent properties is contingent, then the only adequate explanation for the link and the appearance of an emergent property is God's direct activity and stable intention that things be so.[55] Now it seems to many, perhaps most philosophers who do not start with a prior commitment to strong physicalism, that the link is indeed contingent, as seen in several well-known thought experiments.

2.4.5.3 Emergence *Ex Nihilo*

Third, emergence seems to be a case of getting something from nothing, a case of magic without a Magician. If matter is relevantly similar to what current physics and chemistry tell us, and it is bereft of any sort of mental entities, then matter does not have mental potentialities. Given this, it is hard to see how the mere spatial rearrangement of, say, atomic simples to form a more complicated spatial structure would be an adequate cause for bringing into existence *ex nihilo* a completely new sort of entity. This is the problem mentioned above that is allegedly solved by panpsychism.

---

[54] McGinn, *The mysterious flame*, especially 46–54, 66–76, 95–101. One of us has subjected McGinn's view to a careful analysis and critique. See Moreland, J. P. (2008). *Consciousness and the existence of god* (pp. 95–113). London: Routledge.

[55] See O'Connor, *Persons and causes*, footnote 8, 70–71. One of us has argued that the causal relation between mental and physical states is, in fact, due to God's direct activity and stable intention that things remain a certain way. See Moreland, *Consciousness and the existence of god*.

2.4.5.4 Emergence as Brute Facts

Fourth, this “*ex nihilo*” problem generates another difficulty with emergent properties. To see this, consider the following tensed definition of coming-to-be:[56]

> B_t: “X comes into existence” $=_{\text{Def.}}$ There is a property P which is such that *x* has P, and there is no property Q which is such that *x* had Q.

Stated without tense, we have:

> B_{nt}: “X comes into existence at *t*” $=_{\text{Def.}}$ There is some property P which is such that P*x* at *t*, and for all times *t′* prior to *t* there is no property Q such that Q*x* at *t′*.

Now, consider two contrasting views of the composition of time:

> *Instantism*: Temporal intervals are derived and not fundamental entities. Intervals are wholly grounded in dimensionless instants.
>
> *Intervalism*: There are extended temporal intervals that are fundamental in that they are not wholly grounded in the existence of other entities.

B_t and B_{nt} are meant to capture the idea that coming-to-be takes place at an instant of time and is not an extended temporal process or interval. For example, it is not a case of alteration or elanguescence. Further, while consistent with Instantism, these definitions do not entail it. If, as seems likely, intervals have unextended boundaries at their beginning, then coming-to-be would occur at that boundary.

The important point about B_t and B_{nt} is the fact that coming into existence is not a process that could be governed by something, e.g., a (non-Humean) law of nature. Thus, there is no naturalistic or scientific explanation whatsoever as to why one entity comes into existence at *t* as opposed to any other entity. Under the view being considered, emergence is a case of coming-to-be. As such, it leaves as a bizarre brute fact that whenever matter is at such and such a configuration, the same emergent property—e.g., being painful—always emerges. Besides the fact that the naturalist ontology is now chock full of emergent brute facts—an embarrassing state of affairs, given the naturalist’s favorite trump card of ontological parsimony—there is also the improbable coincidence, bordering on a miracle, that the same emergent property always seems to show up when the same physical conditions obtain.

Strawson has given considerable reflection to our last two points.[57] He notes that when it comes to liquidity or heat (the objective physical phenomena and not our experience of them), we have a good story to tell as to how and why these (structural) properties regularly appear under certain circumstances. But when it comes to genuinely emergent properties—he focuses on experiential phenomenal qualia—due to the nature of emergence, there is, in principle, *no* story to tell.[58] He argues that if it is actually the case that Y is emergent from X, then in the relevant sense, it must be the case that Y is wholly dependent on and only on X such that all the features of Y can be metaphysically intelligibly traced, in principle, back to X. If this is not true, then Y did not emerge solely from X.

---

[56] Chisholm, *On metaphysics*, 56.

[57] Strawson, ibid., 12–21.

[58] When John Searle developed his biological naturalism, he dismissed the problem of (genuinely) emergent properties like those constituting consciousness by making a tight analogy between them and structurally supervenient properties, e.g., solidity, liquidity. This allowed him to find emergent mental properties to be ordinary physical, structural ones and their appearance to be unsurprising. See Searle, *The rediscovery of the mind*, xii, 13–19, 25–28, 32, 56–57, 85–93, 95, 118–124. For a response to Searle, see Moreland, *Consciousness and the existence of God*, chapter three.

Based on this, Strawson rightly asserts that emergence cannot be brute since, given the nature of brute emergence; there is absolutely nothing about the intrinsic or relational nature of X in virtue of which Y emerges. That's simply what it means for emergence to be *brute*. *Brutality rules out nothing*. Thus, Strawson concludes that emergence is a genuine miracle, not only because it is the appearance of something from nothing in virtue of which it appears but given that whenever the "appropriate" physical conditions obtain, there is also the bizarre fact that the very same emergent property appears.

But this is utterly ridiculous and, indeed, incoherent. In principle, there is no reason why, says Strawson, a negative number couldn't emerge from the addition of certain positive numbers. We could not agree more with his summary warning:

> If someone says he chooses to use the word "emergence" in such a way that the notion of brute emergence is not incoherent, I will know that he is a member of the Humpty Dumpty army and be very careful with him.[59]

2.4.5.5 Sorites Problems

Fifth, sorites problems are lurking in the neighborhood. Could this emergent property, say being painful, be instantiated with one less atomic part in the subvenient base when that base is the "right" level of complexity? Surely the answer is yes. How about two fewer atomic parts? And so on. *Sans* ontological vagueness, at some point, the emergentist must say that the subtraction of one small atomic part has a huge, disproportionate metaphysical effect on the disappearance of the emergent property or properties, e.g., all the irreducibly mental properties constituting consciousness.

But how can such a significant metaphysical effect be due to such an insignificant cause? And why was the "right" level of complexity one atomic part greater than the actual level? Further, the very notion of "the right level of complexity" is completely vacuous. Why? What is the right level? Answer: Whatever the level is when the emergent property appears. When does the emergent property appear? When the right level of complexity is reached. There are no independent criteria for identifying the "right level of complexity" besides the appearance of the so-called emergent property.

2.4.5.6 The Combination Problem

Sixth, David Chalmers has raised an additional difficulty with SE, given some form of (especially constitutive) panpsychism or (we add) strict microphysicalism—the Quality Combination Problem (viz., plausibly explaining how microqualities combine to yield macroqualities, e.g., phenomenal and moral ones?) of which one especially difficult version is the Palette Problem (viz., Given the highly likely fact that there is a limited number of microqualities but a large number of macroqualities, how do so many get yielded from so few resources?[60]).

---

[59] Strawson, ibid., 18.

[60] Chalmers, D. J. (2017). The combination problem for panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (pp. 204–206). Oxford: Oxford University Press.

Chalmers considers unfavorably two attempts to resolve the Palette Problem: *small-palette solutions*: somehow, all macroqualities can be generated from just a few microqualities, if we find the right ones with sufficient flexibility and generality; and *large-palette solutions*: the full range of macroqualities are included among the microqualities. Against the former, Chalmers argues that it is pure speculation and still faces the Quality Combination Problem (viz., How do microqualities combine to yield *simple* macroqualities?).

Against the latter, Chalmers argues that there are, in fact, only a small number of causally efficacious microqualities; the rest will be epiphenomenal. To deny this, one would have to deny the simplicity of physics with a small number of properties and laws. Says Chalmers, "Large-palette solutions seem once again to be stuck with a form of epiphenomenalism or radical revision of the fundamental dynamics of the physical world."[61]

2.4.5.7 Consciousness and a Bayesian Argument for Theism

Seventh, Ben Page has advanced a Bayesian argument to the effect that the obtaining of the correct conditions for the emergence of consciousness (and, presumably, value properties) in the actual world is substantially more likely, given theism compared to atheism.[62] Thus, these emergent properties (if there are such) confirm theism over atheism.

The prior probability of conscious (and intrinsically valuable) beings is higher on theism than on atheism. Why think this? Page argues that given theism, God would have several reasons for creating conscious (and intrinsically valuable) beings and bring about whatever conditions that would do the job: It seems likely that God would want to have love relationships with creatures, to share His knowledge and what it is like to experience various things with them; being creative, God desires to create all kinds of beings, those like human persons being one kind, and so on. But in atheistic hypotheses about the actual world, there is no such being who wants to bring about creatures like us. There is no teleological driving force behind the laws of nature that would be directed toward bringing about the conditions required for our coming-to-be.[63] Rather, on atheist worlds with laws like ours, it is extremely unlikely that those conditions would obtain.

To make this more explicit, Robert Koons has argued that all else being equal, emergence is an extremely complicated and unlikely alternative to a simple soul.

> We should avoid brute conspiracies among large numbers of independent agents. Emergence would require a very improbable and *ad hoc* preestablished harmony among the powers of the many mutually unifying parts—a coordinated distribution of mutually exercisable powers. This problem ramifies as the number of

---

[61] Chalmers (2017), 206.

[62] Page, B. (2020). Arguing to theism from consciousness. *Faith and Philosophy*, *37*(3), 336–362.

[63] Thomas Nagel sees the problem here and advances a non-theistic teleological view of our universe to overcome it. See *Mind and cosmos*. (2012). Oxford: Oxford University Press. For a critique, see Moreland, J. P. (2012). A reluctant traveler's guide for slouching towards theism: A philosophical note on Nagel's *Mind and cosmos*. *Philosophia Christi*, *14*, 415–424.

> components to be unified increases. It becomes quite untenable when billions of components must unite with each other.[64]

2.4.5.8 Novel Properties and New Substances

Eighth, since our concern in this book is with issues relevant to the reality of a simple self or soul, in offering this final objection, we limit our focus to that range of properties that constitute consciousness, e.g., being such and such a sensation, being a thought that p.

Throughout the history of philosophy from Aristotle to the rise of the atomistic mechanical philosophy in the scientific revolution of the 17th century, the appearance of a *sui generis* new kind of property was taken to be an indication of a new classic substance (with notable exceptions, construed along Aristotelian lines) that exemplified it. If we take an emergent property to be, minimally, a *sui generis* new kind of property *that wholly comes from, is rooted in, and sustained by and only by base entities*, there were no such things as emergent properties. New kinds of properties usually indicated a substantial change had occurred and, since substances were construed as fundamental, the new properties were taken to be attributes of the new substance grounded in its substantial form.[65]

With precursors in Galileo and a few others, the 17th century witnessed a pervasive adoption of Democritean atoms or corpuscles and a view of chemical change as accidental change involving the rearrangement of unchanging atoms. Real generation and corruption were replaced with alteration. As time went on, though changes took place regarding the original conception of these atoms (e.g., the shift from inert to dynamic corpuscles), the basic picture continued to flourish.[66] In light of this viewpoint, any appearance of a genuinely new kind of property had to be taken as brute emergence.

By way of application, there are two ways to incorporate the irreducible properties that constitute consciousness: they are either emergent properties or properties exemplified by a new substance, a soul. We believe the choice here is a metaphysical and not a scientific one. We also find the widespread employment of emergence in this context to be question-begging, usually without so much as a nod toward the alternative view. One of the goals of this book is to provide adequate grounds for believing in a simple, substantial soul. If this is correct, then the properties constituting consciousness are not emergent; they are properties exemplified by a specific sort of new substance, viz., a soul.[67]

---

[64] Koons, R. C. (2018). Against emergent individualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 377–393). Oxford: Wiley-Blackwell. The quote is on pages 384–385. In a related context, as a part of defending his version of noncausal libertarian agency, E. J. Lowe presents this argument. He claims that even if such an event (e.g., raising one's hand to vote) is one for which we have identified all the physical causes relevant to it, this does not preclude the explanatory importance of an exclusively nonphysical state (e.g., an intention) in accounting for why the relevant body movement type, which has myriads of alternative combinations of physical causal pathways, occurred in a noncoincidental way, an explanation that the mere tracing of a variety of causal pathways cannot explain. See Lowe, E. J. (2008). *Personal agency* (pp. 33–34). Oxford: Oxford University Press.

[65] Though we do not agree with her overall position or her faint-hearted hylomorphism, and we think she misses the dependency aspect of emergence *vis a vis* substantial change, nevertheless, Eleonore Stump has provided a helpful comparison between emergent properties and new properties resulting from substantial change. See her (2013). Emergence, causal powers and Aristotelianism in metaphysics. In R. Groff & J. Greco (Eds.), *Powers and capacities: The New Aristotelianism* (pp. 48–68). New York, NY: Routledge.

[66] See Yolton, J. W. (1983). *Thinking matter: Materialism in eighteenth-century Britain*. Minneapolis, MN: The University of Minnesota Press.

As Uwe Meixner notes, it is hard to see how emergentism is a thesis of physicalism. Physicalism, says Meixner, is a subsumptive thesis, viz., all nonabstract individuals are subsumed under the completely physical ones.[68]

In conclusion, while Timothy O'Connor is an expert on and advocate of emergent properties, nevertheless, he admits that as science has progressed, it is harder to find genuine cases of them, and the number of plausible cases is shrinking.[69] Moreover, O'Connor makes a telling observation:

> It bears emphasis that strong emergentists typically also suppose that some interesting organized behavior is more plausibly understood either reductively or as weakly emergent: their strong emergentism is piecemeal.[70]

The only plausible example of emergence that O'Connor can find are properties constitutive of phenomenal consciousness. In a similar vein, Mark Bedau warns against adopting emergent properties (he calls them strong emergent properties):

> All the evidence today suggests that strong emergence is scientifically irrelevant ... We should accept brute natural phenomena only if we are convinced that they cannot be explained. The problem is that strong emergence seems like it can always be explained away.[71]

And for Bedau, the only plausible candidate for strong emergent properties are mental properties. In the chapters to follow, we will argue in various ways that irreducible phenomenal properties of consciousness are not emergent from some underlying base; rather, they are properties of a new simple mental substance.

2.5 GOING FORWARD

We are convinced that worldview considerations, especially naturalistic ones, play important ontological and epistemological roles in topics in philosophy of mind and conversely. In particular, staunch naturalism/strong physicalism set the ontological and epistemological parameters for a large number of contemporary philosophers of mind. Moreover, in this chapter, we have offered a critique of genuinely *sui generis* emergent properties and, along with it, of faint-hearted naturalism and weak (emergent) physicalism. Thus, in our view, a naturalist *ought* to be a staunch naturalist and strong physicalist. Many of the arguments to follow stand alone and are not explicitly related to these background worldview issues. But many of them are explicitly related, and, where appropriate, we employ worldview considerations in stating and evaluating various arguments. Even in cases where background worldview considerations are not explicitly involved, those considerations are often implicitly involved, and making explicit what is implicit will be part of the dialectic when needed. We hope that this worldview approach will be illuminating, refreshing, unique, and profitable to the reader.

---

[67] C. B. Martin has presented a spirited defense of an idiosyncratic view of emergent properties. See Martin, C. B. (2007). *The mind in nature* (pp. 129–139). Oxford: Clarendon. Cf. 1–11, 35–35–53. For evaluations of Martin's views, see Jessica Wilson's review of Martin's book in *Mind* 119 (April 2010): 503–511; and Matheson, C. (2009). Review of *The mind in nature*. *History of Philosophy and Logical Analysis*, *12*(1), 315–320.

[68] Meixner, U. (2016). Against physicalism. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 17–34). London: Routledge. The relevant material is found on pages 17–19.

[69] O'Connor, T. *Persons and causes* (pp. 110–121).

[70] O'Connor, T. (2021, Winter). Emergent properties. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (2021 ed.). https://plato.stanford.edu/archives/win2021/entries/properties-emergent.

[71] Bedau, p. 51.

CHAPTER 3

A Brief Ontological Detour: Subjects, Parts, Wholes, and Unity

Among his many insights into the state of contemporary philosophy, E. J. Lowe rightly observed that "… much of the so-called 'metaphysics' in contemporary philosophy of mind is really rather lightweight, often having only a tenuous relation to serious foundational work in ontology."[1] Steven Horst makes a similar point about how contemporary philosophy of mind is blind to contemporary philosophy of science.[2] As an attempt to make the progress urged by Lowe, we will explore a detailed mereology, which we will employ in arguments for substance dualism in various remaining chapters. We also develop our understanding of the metaphysical relation between subjects of phenomenal consciousness and their phenomenal states, which will be employed in later chapters.

3.1 HUSSERLIAN PARTS, WHOLES, AND UNITY

Husserl's ontology combines an Aristotelian constituent ontology according to which genuine universals—properties and relations (propositions are structural mental properties)—are immanent within and constitute the species/nature of their metaphysically complex instances (or concrete particulars) with a rich mereological ontology.[3]

---

[1] Lowe, E. J. (2000). *An introduction to the philosophy of mind* (p. xi). Cambridge: Cambridge University Press.

[2] Horst, S. (2007). *Beyond reduction: Philosophy of mind and post-reductionist philosophy of science* (pp. 23–64). Oxford: Oxford University Press.

[3] See Moreland, J. P. (2013). Exemplification and constituent realism: A clarification and modest defense. *Axiomathes*, *23*, 247–259.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

3.1.1 Parthood and Simplicity: Separable and Inseparable Parts

Many consider Husserl's third *Logical Investigation* as, in the words of Kit Fine, "perhaps the most significant treatise on the concept of part to be found in the philosophical literature."[4] Husserl makes several part–whole distinctions, including kinds of dependency relations among an object's parts, a hierarchy among parts involving a detailed conception of levels, and a theory involving various kinds of unity an object can possess.[5] Like Aristotle and Franz Brentano, Husserl is a parthood pluralist.[6] For Husserl, "part" broadly refers to any distinguishable, particularized aspect of an object, a particular "in" a whole.

Husserl specifies two kinds of parts: "pieces" and "moments." *Pieces* have a degree of metaphysical independence or separability from the wholes of which they are a part. Conversely, *moments* are "non-independent" or inseparable from the wholes of which they are a part. Husserl's pieces and moments are sometimes called dependent or inseparable parts and independent or inseparable.[7] We will use "inseparable part" in place of "moment" and "separable part" in place of "piece."[8]

Separable and inseparable parts differ concerning ontological dependence. *Separable parts* are independent in that they can retain their identity and exist without the whole of which they are a part. Such parts are accidental to other such parts and their wholes in that they require an external relation to bond together into their wholes (e.g., mereological aggregates of separable parts). Such bonding does not follow from the nature of the parts.

Conversely, *Inseparable parts* depend on the whole of which they are a part for both their identity and existence (metaphysically and conceptually). They are intrinsically partial in this sense: they require supplementation by something else to exist. That supplemental entity is the inseparable part's foundation. Foundation can be one-sided (e.g., an instance of redness cannot exist without foundation in an instance of extension but not conversely) or mutual (an instance of size cannot exist without an instance of shape and conversely.) The essence of the part makes it separable, not the mere ability to conceive of it apart from its whole.[9]

Inseparable parts cannot be conceived of nor exist as an independent whole separate from the whole of which it is a part. Inseparable parts stand in the modal distinction to their wholes and, in this sense, may be properly understood as modes of their wholes. Husserl offers various examples of inseparable parts. An instance of color, for example, cannot be presented—exist or be conceived of—apart from an instance of extension (e.g., apart from the surface on which it is located), and is inseparable from the object of which it is a part.[10] Instances of extension, surface, color, and brightness are all inseparable parts of a material object taken as a whole. An inseperable part is not, explains Husserl, "merely bound up with its associates, but blends with them." The inseparable parts of a whole permeate each other and the whole of which they are parts.[11]

---

[4] Fine, K. (1995). Part-whole. In B. Smith & D. W. Smith (Eds.), *The Cambridge companion to Husserl* (p. 463). Cambridge: Cambridge University Press.

[5] Husserl, E. (1970). *Logical investigations* (J. N. Findlay, Trans., Vol. 2, pp. 466–481). London: Routledge and Kegan Paul. (First published in German, 1900–01).

[6] See, e.g., Chisholm, R. M. (1986). *Brentano and intrinsic value*. Cambridge: Cambridge University Press; and Moreland, J. P. (2002). Naturalism, nominalism, and husserlian moments. *The Modern Schoolman, 79*, 199–216.

[7] See, e.g., Willard, D. (1994). Mereological essentialism restricted. *Axiomathes*, *5*(1), 123–144. The same is true of the mereology of Husserl's mentor, Franz Brentano. Cf., Salice, A. (2017). Franz Brentano. In J. Seibt, G. Imaguire, & S. Gerogiorgakis (Eds.), *Handbook of mereology* (pp. 119–125). München, Germany: Philosophia Verlag; and Mulligan, K., & Smith, B. (1985). Franz Brentano on the ontology of mind. *Philosophy and Phenomenological Research*, *45*(4), 627–644.

[8] This section is heavily informed by the following. Rosiak, M. (2017). Edmund Husserl. In J. Seibt, G. Imaguire, & S. Gerogiorgakis (Eds.), *Handbook of mereology* (pp. 262–269). München, Germany: Philosophia Verlag; McCarthy, J. C. (1992). Parts' whole, and the forms of life: Husserl and the new biology. In L. Hardy & L. Embree (Eds.), *Phenomenology of natural science* (pp. 135–156). Netherlands: Kluwer Academic Publishers; and Sokolowski, R. (1968). The logic of parts and wholes in Husserl's investigations. *Philosophy and Phenomenological Research*, *28*(4), 537–553.

[9] Husserl, §7, pp. 446–447.

Separable parts, on the other hand, are separated from the other parts of the whole. Separable parts are "what they are no matter what goes on around them,"[12] even if "everything outside it were annihilated."[13] Separable parts are, as it were, scattered throughout the whole of which they are parts. We can now state necessary and sufficient conditions for separable and inseparable parts as follows.

> *Separable Part*: *x* is a separable part of some whole *W iff* (i) *x* is a particular, (ii) *x* is a part of *W*, and (iii) it is possible for *x* to exist without being a part of *W*.
>
> *Inseparable Part*: *x* is an inseparable part of some whole *W iff* (i) *x* is a particular, (ii) *x* is a part of *W*, and (iii) it is not possible for *x* to exist without being a part of *W*.

The paradigm case of an inseparable part in this tradition of Brentano and Husserl is a monadic property-instance. If substance s has property P, <the-having-of-P-by-s> is (i) a property-instance of P; and (ii) an inseparable part of s, what we also call a mode of s. Assuming for the sake of argument that a lump of clay is a substance (most likely, it is an ordered aggregate, not a substance) and that it has a spherical shape, then the property <being spherical> is a universal attribute, and <the-having-of-sphericity-by-the-clay> is a mode (inseparable part) of the clay and a property-instance of sphericity.

Setting aside properties, there are two ways something can be simple in the sense relevant to our discussion: by not having any separable parts or by being metaphysically indivisible. We use "metaphysically indivisible" to mean what many philosophers say by "indivisible in thought." Something could be metaphysically divisible but not physically divisible (if, say, such division annihilated the whole), but not conversely. Moreover, all metaphysically indivisible particulars are uncomposed (e.g., Markosian's pointy objects), but not conversely.[14] A continuously extended whole with no separable parts could be divided.[15]

According to our usage, which is compatible with Markosian's *Pointy View*, a substance with inseparable parts is mereologically simple. More specifically,

> *Mereological Simplicity*: *x* is mereologically simple $=_{\text{def.}}$ *x* has no separable parts.

---

[10] Husserl (1900/1970), §4, p. 440.

[11] Husserl (1900/1970), §8, p. 449.

[12] Husserl (1900/1970).

[13] Husserl (1900/1970), §6, p. 445.

[14] *The Pointy View*: *x* is a mereological simple $=_{\text{def.}}$ *x* is point-sized (*x* may have spatial location but not extension). This may, but need not, be combined with a pointy view of space in which case a simple entity overlaps with and only with an unextended spatial point. See Markosian, N. (1998). Simples. *Australasian Journal of Philosophy*, *76*, 332–340.

[15] In Chapter 7 we explore this in more detail.

Notice, the fact that *x* is mereological simple is compatible with *x* having Husserlian *inseparable parts*.

Notice also that the various extant answers to the special composition question regarding separable parts and the (so-called) wholes they compose, especially the bond state and the joining-relation views, fail to result in a whole that is a genuine classic (Aristotelian) substance. Why? Because the unity of that whole is too weak for it to count as having substantial unity. Over several centuries, to secure that unity, the fundamental axiom of the part–whole relation for substances has been this: No substance has other substances as proper parts. In Husserlian language, this becomes "No substance has separable parts as proper parts." We will elaborate and apply these insights now, and to the unity of consciousness in Chapter 6.

Any relevant entity to which I am identical on a physicalist view is (most plausibly) composed of separable parts (the body, brain, or a standard subregion of the brain—understand according to a physicalist ontology) or (less likely) at least divisible (a physical simple if such there be). It might be objected, however, that in the history of science, there have been coherent conceptions of matter according to which matter is not extended. We are thinking of the penetrable, unextended point particles of Roger Joseph Boscovich. Critics of Boscovich claimed that his position amounted to "the immateriality of matter," and that his point particles fit more naturally into a Berkeleyan ontology than a materialist one.

3.1.2 Holistic Unity: Genuine Wholes, Relations, and Aggregates

Various relations obtain between the parts of a whole and between the parts and the whole itself. As noted, Husserl was a realist regarding universals. Besides monadic properties, he held to the existence of what we would call external and internal relations. Roughly, internal relations stand "between elements within a given whole," while external relations stand "between that whole (or elements in that whole) and something outside the whole."[16] Regarding internal relations, Gustav Bergmann explains that "*The ontological ground of an internal connection is the natures of the entities it connects and nothing else*."[17]

Internal relations also possess modal characteristics not possessed by external relations. According to D. M. Armstrong, *R* is an internal relation $=_{\text{def.}}$ *R*'s holding between *a* and *b* is necessitated by the intrinsic natures of *a* and *b*.[18] Internal relations are necessitated by their relata, such that if *a* and *b* are inseparable parts of some whole *W*, then there is no possible world in which *W* exists and where *R* does not obtain between *a* and *b*. The nature of inseparable parts, as such, exemplifies internal relations, while the nature of separable parts, as such, exemplifies *only* external relations.

We understand internal relations as follows.

> *Internal Relation*: *R* is an internal relation between *a* and *b iff* (i) facts about *R* are grounded in facts about the natures of *a* and *b*, and (ii) necessarily, if *R* fails to obtain, *a* and *b* are altered.

---

[16] Ewing, A. C. (1934). *Idealism: A critical survey* (p. 118). London: Methuen.

[17] Bergmann, G. (2004). Realism: A critique of Brentano and Meinong. In E. Tegtmeier (Ed. and Intro.), *Collected works volume 3* (p. 54). Frankfurt: Ontos Verlag. First published, 1967.

[18] See, e.g., Armstrong, D. M. (1978). *A theory of universals*, vol. 2 of *Universals and scientific realism* (pp. 84–85). Cambridge: Cambridge University Press; Armstrong, D. M. (1989). *Universals: An opinionated introduction* (pp. 43–44, 55, 100). Boulder, CO: Westview Press; and Armstrong, D. M. (1997). *A world of states of affairs* (pp. 87–89). Cambridge: Cambridge University Press.

If the *R* of *aRb* is internal to *a* (or both *a* and *b*), then for all *x*, if *x* does not stand in *R* to *b*, then $x \neq a$. An internal relation is grounded in the nature/essence of the relevant relatum or relata. Conversely, we understand external relations as,

> *External Relation*: *R* is an external relation between *a* and *b iff* (i) facts about *R* are not grounded in facts about the natures of *a* and *b*; and (ii) possibly *R* fails to hold between *a* and *b*, while at least *a* or *b* remains unaltered.

If the relation, *R*, of *aRb* is *external to a* (or *a* and *b*) then *a* can exist and retain identity if it fails to stand in *R* to *b*. Spatial relations, for example, are external.

Plausibly, every pair of objects has at least one internal relation, that of co-parthood. It, therefore, follows that for every pair *a*-and-*b*, the objects of that pair, *a* and *b*, will stand in at least one internal relation to each other, although some pairs also have external relations.

As we will see shortly, this distinction, along with Husserl's separable/inseparable part distinction, determines the natures of and distinction between a *genuine whole* and an *aggregate*, and between *holistic unity* and *functional unity*.

Consider first a distinction between a genuine whole and an aggregate *taken as a whole*. According to Husserl, it is in the essence of an inseparable part that it is united (by internal relations) with its correlated whole. Such wholes need no principle of unity. Husserl refers to these wholes as genuine wholes.

> *Genuine Whole*: *W* is a genuine whole *iff* (i) *W* is a whole, (ii) *W* has only inseparable parts, and (iii) *W* is unified intrinsically by, and only by, the nature of *W*'s inseparable parts.

The same is not true of wholes (understood in a very weak sense) with at least one separable part. Husserl calls such wholes aggregates.

> *Aggregate*: An object *O* is an aggregate *iff* (i) *O* has at least two separable parts, and (ii) the separable parts of *O* stand only in external relations between each other and *O*.

Consequently, necessarily, there is no *O* such that *O* is both a genuine whole and an aggregate. If *O* has both separable and inseparable parts, *O* cannot be a genuine whole, as genuine wholes do not have separable parts. Even if the inseparable parts of *O* could be unified intrinsically, it would not follow that *O* as a whole is intrinsically unified, as *O* also has separable parts that stand only in external relations to each other. Consequently, for all *O*, if *O* is a whole, then *O* will be either a genuine whole or an aggregate.[19] The logical space is exhausted by these options. According to Husserl, for any genuine whole *W*, *W* is unified by internal relations between the inseparable parts of *W* and *W* itself.

---

[19] Here is a possible concern. Supposing mereological universalism is true, then everything that exists has a fusion. Suppose that numbers are necessary objects. This means that the numbers 1, 2, 3 form a fusion, which exists necessarily, the fusion [1, 2, 3]. This fusion turns out to be a genuine whole. Here is the worry: this counterexample shows that Genuine Whole, as we've stated it, does not capture the concept of a genuine whole, because intuitively fusions are not genuine wholes. This does not strike me as a serious concern for the following reasons. First, it isn't clear to us that mereological fusions are genuine wholes. It seems the members of a fusion are separable parts as they can exist and retain their identity outside the fusion. Alternatively, we could understand this counterexample as showing that our intuitions about fusions are mistaken, at least concerning a minimal set of fusions, those with parts that exist necessarily. To accommodate this concern, Genuine Whole can be stated disjunctively: *x* is a genuine whole *iff* (i) *x* is a whole, and (ii) *x* is either a mereological fusion, where the parts of such a fusion exist necessarily or (i) *W* is a whole, (ii) *W* has only inseparable parts, and (iii) *W* is unified intrinsically by, and only by, the nature of *W*'s inseparable parts.

The nature of a genuine whole and an aggregate also determine the kind of unity each has. Genuine wholes possess what Leibniz called "true unity" or what we will call holistic unity. In contrast, aggregates may possess a much weaker functional unity. We understand this distinction as follows.

> *Teleological Holistic Unity*: *O* is holistically unified *iff* (i) *O* is a genuine whole, and (ii) the parts of *O* work together to accomplish some end as a consequence of the relations among *O*'s parts.

As we will explain shortly, this teleological characterization of holistic unity, although the dominant traditional view, has fallen out of style. Here is a naturalist-friendly understanding of holistic unity.

> *Holistic Unity*: *O* is holistically unified *iff O* (i) *O* is a Genuine Whole, and (ii) the parts of *O* cause some effect as a consequence of the internal relations among *O*'s parts.

We will employ Holistic Unity throughout the following chapters as our arguments do not require anything more. In relation to Teleological Holistic Unity, one might adopt the following traditional thesis.

> *Teleological Functional Unity*: *O* is teleologically functionally unified $=_{\text{def.}}$ (i) *O* is an aggregate, and (ii) the separable parts of *O* work together to accomplish some end as a consequence of the relations (internal or external) among *O*'s parts.

Like Teleological Holistic Unity, we recognize that this definition incorporates teleological language that most would want to reduce to some etiological expression. Therefore, one might define functional unity as follows: *O* is functionally unified $=_{\text{def.}}$ (i) *O* is an aggregate, and (ii) the separable parts of *O* work together to accomplish some end as a consequence of the external relations among *O*'s parts. At first glance, this definition seems to adequately capture what one is getting at by using "functional unity." But as it stands, it is not adequate. For one thing, the canonical view of function over the last several centuries has characterized it without employing teleological language, especially for mereological aggregates. Some version of *etiological reduction* is usually used to provide a reductive analysis of teleology. It is not our purpose to explore those reductions here. Instead, we will offer a non-teleological definition of "functional unity" that leaves open different ways to fill it out, e.g., with some sort of evolutionary chain efficiently causing some reproductively advantageous effect. More specifically,

> *Functional Unity*: *O* is functionally unified $=_{\text{def.}}$ (i) *O* is an aggregate, and (ii) the separable parts of *O* work together to cause some effect as a consequence of the external relations among *O*'s parts.

Although this definition is specifically an etiological functional unity thesis, we will use the name Functional Unity for simplicity and because our arguments will run through on this naturalist-friendly functional unity thesis.

One way to understand the unity of a Genuine Whole is Husserl's view that necessity "is packed into" dependence and dependence into inseparable parts. For Husserl, *a* is dependent, or founded, on *b*, just in case, necessarily, *a* exists only if *b* exists. This is not, however, a brute fact on Husserl's view. As Willard explains, "the existential dependence (or lack thereof) is derivative from the essence or nature of the contents concerned, from the properties constitutive of what they are."[20]

These dependency relations are governed by what Husserl calls "laws of essence." If *a* depends on *b*, it will be in virtue of the essences of *a* and *b*. The fact that *a* is an inseparable part of *b* is grounded in facts about the nature of *a* and *b*. Consequently, the dependency of the inseparable parts of a whole *W* is a result of the internal relations, the laws of essence, among both the inseparable parts of *W* as well as among the inseparable parts and *W* itself.[21] The part–whole relation consists of the inherent connections and disconnections between properties (or essences) that constitute both the natures of the wholes and the parts of the whole.[22] Husserl's view is, therefore, a kind of priority holism or substantial priority thesis. A Genuine Whole, *W*, is prior to *W*'s inseparable parts. The inseparable parts of *W* receive their identity from how they are internally related to *W*.

As Figure 3.1 illustrates, the holistic unity of an entity is a product of its internal nature; it resides within the entity. The internal relations of an entity do the work of a principle of unity. Conversely, an aggregate possesses only functional unity, a unity that is not the product of anything within the aggregate itself but of the external relations among its separable parts. Accordingly, the internal structure of a Genuine Whole includes a set of internal relations among its inseparable parts, which provide for the holistic unity of the genuine whole. Similarly, the structure of an aggregate, the set of the aggregate's external relations, determines the functional unity of the aggregate. Therefore, only mereologically simple entities, substances without separable parts, are holistically unified, as true unity is grounded in the kind of dependence or "foundation" that exists uniquely between wholes and their inseparable parts.[23]

In summary, the chain of reasoning we've used in this section moves from facts about mereologically simple objects to holistic unity and facts about complex objects to functional unity. Figure 3.2 is helpful.

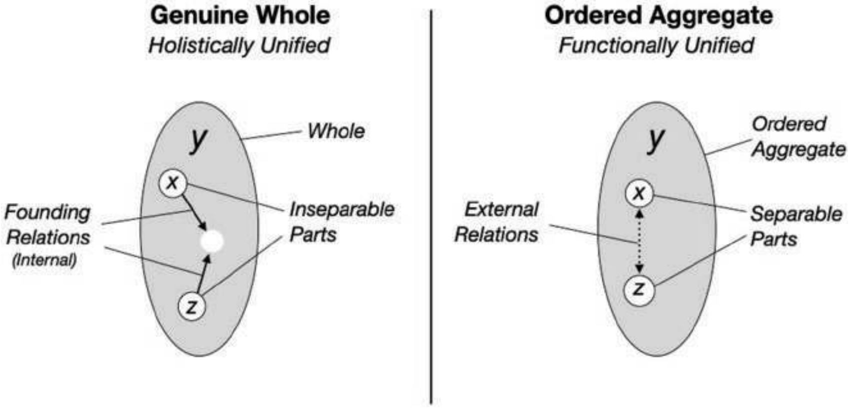


FIGURE 3.1 Genuine Wholes and Ordered Aggregates.

---

[20] Willard, "Mereological Essentialism Restricted," 141.

[21] Husserl, *Logical investigations*, §22–23, 476–481.

[22] Willard, 142.

[23] Husserl, §22, 478.

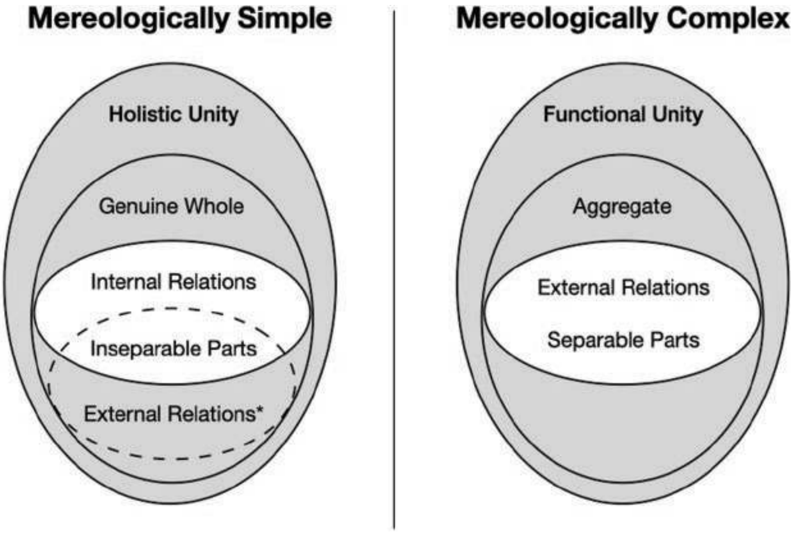


FIGURE 3.2 Mereological Simplicity and Complexity.

This figure is not intended to chart the logical space. It illustrates that (i) if *x* is holistically unified, then *x* is a genuine whole having only inseparable parts which stand in only internal relations to each other; and (ii) if *x* is functionally unified, then *x* is an aggregate with at least one separable part standing in external relations to each other.[24]

3.2 SUBJECTS OF CONSCIOUSNESS: A FEW METAPHYSICAL THESES

3.2.1 What Is Consciousness?

Consciousness is familiar and mysterious. It is part of every experience we have, but it is difficult to define what we mean by the term "consciousness" in much detail. We suggest, as do others, that we should first define "consciousness" by ostension. We point to or exemplify the various kinds of experiences and their commonalities. From there we refine our definition of "consciousness." Consider the following uncontroversial instance of regaining consciousness as if it were your own.

> *Coma*: Suppose I slowly come out of this comatose state. I experience my body resting against the hospital bed, warm and relaxed, but with some stiffness. I soon hear a family member talking to a doctor. I feel the dryness of my mouth and have a desire for water. As I open my eyes, I experience the brightness of hospital lights. I have an experience of grogginess and feel it slowly fade as my experiences and thinking become clearer.

I go from unconscious to conscious, slowly regaining various sensations, thoughts, perceptions, and desires. But, of course, this is only where we start. We revise our conception of consciousness, as we do with most concepts, as we get more evidence. What seems indispensable to our set of evidence is the phenomenological data available in our own experience from the first-person perspective.

---

[24] This construal makes it possible that a mereologically simple object has inseparable parts that are unified by both internal and external relations. Such an object would be holistically unified by the internal relations, and unified in some way addition way, which does not require having separable parts.

Some might be suspicious of defining consciousness by ostension. I don't see a problem. First, what counts as physical is initially defined by ostension. Second, plausibly, all knowledge depends on what we can only define or understand by ostension. As Bertrand Russell observes,

> all nominal definitions, if pushed back far enough, must lead ultimately to terms having only ostensive definitions, and in the case of an empirical science the empirical terms must depend upon terms of which the ostensive definition is given in perception.[25]

"All our knowledge," says Russell, "both knowledge of things and knowledge of truths, rests upon acquaintance as its foundation."[26] All knowledge ultimately depends on our direct acquaintance with reality, including consciousness. This will be explored in Chapter 5.

This approach might seem misguided to those suspicious of knowledge by acquaintance and convinced that consciousness should be defined exclusively by neuroscience.[27] However, defining consciousness neuroscientifically is problematic for a variety of reasons. First, defining consciousness exclusively by neuroscience immediately violates multiple realizability, the thesis that a given psychological kind (like pain) can be realized by many distinct physical kinds (e.g., by brain states of various kinds of biological beings, or by electronic states of computers).

Second, neuroscience is flooded with competing definitions of consciousness,[28] and there is scarce agreement on how to approach a definition.[29] Nor are there agreed-upon measures for identifying consciousness at the behavioral level. Third, it is too easy to conflate conscious experience with cognitive functions that are only necessary for the subject to report having a conscious experience. As a result, it is extremely difficult to evaluate independent neurophysiological claims.[30] As Eric Schwitzgebel concludes,

> one can either be a purist and do without consciousness (or operationalize the term behaviouristically) or one can do one's best to muddle forward through doubt and ambiguity.[31]

---

[25] Russell, B. (1948). *Human knowledge: Its scope and limits* (p. 242). New York: Simon and Schuster.

[26] Russell, B. (2001). *Problems of philosophy* (2nd ed., p. 26). Oxford: Oxford University Press. First published, 1912.

[27] Contemporary philosophy has focused primarily on propositional knowledge. However, knowledge by acquaintance is making a significant comeback. See, e.g., Raleigh, T., & Knowles, J. (Eds.). (2019). *Acquaintance: New essays*. Oxford: Oxford University Press.

[28] See. e.g., the 2009 special edition of the *Journal of Consciousness Studies* on defining consciousness, edited by Chris Nunn, especially the following paper: Vimal, R. L. P. (2009). Meanings attributed to the term "consciousness": An overview. *Journal of Consciousness Studies, 16*(5): 9–27.

[29] See, e.g., Pereira Jr., A., & Ricke, H. (2009). What is consciousness? Towards a preliminary definition. *Journal of Consciousness Studies*, *16*, 28–45.

[30] See, e.g., Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11): 494–501.

[31] Schwitzgebel, E. (2007). Do you have constant tactile experience of your feet in your shoes? Or is experience limited to what's in attention? *Journal of Consciousness Studies*, *14*, 32.

It is not uncommon for neuroscientists to make admissions of the following sort:

> Either the EEG [a tool used to register brain activity] is not sensitive enough to the neural processes underlying consciousness, or we still do not yet fully understand what to look for.[32]

Peter Bandettini, regarding the limits of fMRI data interpretation, warns against the far too common problem of "drawing unsubstantiated inferences about mental states."[33] A recent meta-analysis found that commonly used task-fMRI measures are generally unreliable for biomarker discovery and brain–behavior mapping.[34] After asking how conscious experience arises from physical systems, Steven Pinker answers, "Beats the heck out of me. I have some prejudices, but no idea of how to begin to look for a defensible answer. And neither does anyone else."[35] After surveying the recent history of neuroscience, Donald Hoffman observes, "In short, the scientific study of consciousness is in the embarrassing position of having no scientific theory of consciousness."[36] The capabilities of neuroscientific research regarding the nature of consciousness are far more inadequate than most realize. Defining consciousness by ostension does not eschew neuroscience but recognizes its limitations. Some neuroscientists acknowledge this. As Gregg Rosenberg states,

> We identify phenomenal consciousness by being acquainted with it, not by looking up a scientific definition ... If we need a definition, the best we can do is to create an operational definition by calling attention to it in increasing levels of detail.[37]

We will take a first-person approach to analyzing consciousness,[38] an approach that takes subjective experience as a source of indispensable, albeit defeasible evidence.[39] Like others, we are suspicious of the authority given to third-personal theorizing about consciousness.[40] It isn't clear to us that there is such a thing as a third-person perspective, but rather a third-person mode of description.[41] Moreover, it is from the first-person perspective that we have conscious experience.

---

[32] Alkire, M., et al. (2008). Consciousness and anesthesia. *Science*, *322*(5903), 877.

[33] Bandettini, P. A. (2009). What's new in neuroimaging methods? *The Year in Cognitive Neuroscience, Annals of the New York Academy of Sciences*, *11*(56), 262.

[34] A meta-analysis of 90 experiments (N = 1,008) revealed poor overall reliability—mean intraclass correlation coefficient (ICC) = .397. The test-retest reliabilities of activity in a priori regions of interest across 11 common fMRI tasks collected by the Human Connectome Project (N = 45) and the Dunedin Study (N = 20) were poor (ICCs = .067–.485). Elliot, M. L., et al. (2020). What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychological Science*, *31*(7), 792–806.

[35] Pinker, S. (1997). *How the mind works* (pp. 146–147). New York, NY: W. W. Norton.

[36] Hoffman, D. D. (2008). Conscious realism and the mind-body problem. *Mind and Matter*, *6*(1), 90.

[37] Rosenberg, G. H. (2004). *A place for consciousness: Probing the deep structure of the natural world* (pp. 3–4). Oxford: Oxford University Press.

[38] Although a first-person approach is not especially popular among naturalists, it is implicit in standard neuroscience experiments and explicit in some approaches to neuroscience in general. See, e.g., Manzotti, R., & Moderato, P. (2014). Neuroscience: Dualism in disguise. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 81–87). New York, NY: Routledge.

[39] Unless specified otherwise, terms such as "consciousness" and its cognates, and phrases such as "being conscious" should be understood as referring to phenomenal consciousness. Furthermore, as explained in Chapter 1, we reject the common conception of a mental state as conscious. By a "state of phenomenal consciousness," we mean the mental state of a subject in virtue of which there is something-it-is-like for the subject to be in that state. Consequently, we will employ this view when stating standard accounts of certain features of consciousness.

3.2.2 Realism About Substantial Subjects and the Self

We are committed to realism about persons or selves. Contemporary notions of the self are muddled. In one view, the self is a product of modernity, a subjectively constructed conception of one's psychological constitution (including one's ethical concerns and narrative self-understanding).[42] This is not what we mean by the self. In fact, the claim that the self is a modern creation is demonstrably false.[43] The self, metaphysically understood, is fairly commonsensical. Galen Strawson explains,

> I propose that it is (at least) the sense that people have of themselves as being, specifically, a mental presence; a mental someone; a single mental thing that is a conscious subject of experience, that has a certain character or personality, and that is in some sense distinct from all its particular experiences, thoughts, and so on, and indeed from all other things.[44]

The self then is at least a metaphysical constituent of reality, a subject of consciousness, a unified substance, and that to which the indexical "I" refers. As Hume and others urge, the self is not eliminable, a mere illusion. The self is fundamental, irreducible. The self is not a mere "Center of Narrative Gravity,"[45] "collections of experiential powers,"[46] or a "virtual center of phenomenal gravity."[47] Nor is the self an incomprehensible mystery,[48] a "grammatical illusion,"[49] or "a concept that does not exist."[50] By a self or subject of consciousness, we have in mind a weaker form of E. J. Lowe's notion of a person as a *psychological substance*, where "a person is a substantial individual belonging to a natural kind which is the subject of consciousness ..."[51] Thomas Reid famously articulated this view.

---

[40] Siewert, C. (1998). *The significance of consciousness*. Princeton, NJ: Princeton University Press, chapter 2; and Meixner, U. (2004). *The two sides of being: A reassessment of psycho-physical dualism* (pp. 270–275). Paderborn: Mentis-Verlag.

[41] See, e.g., Meixner, *The two sides of being*, 270–275.

[42] This thesis was made popular by Richard Taylor. See, Taylor, C. (1989). *Sources of the self: The making of the modern identity*. Cambridge: Cambridge University Press.

[43] See, e.g., Martin, R., & Barresi, J. (2006). *The rise and fall of soul and self: An intellectual history of personal identity*. New York, NY: Columbia University Press.

[44] Strawson, G. (1997). The self. *Journal of Consciousness Studies*, *4*(5–6), 407.

[45] Dennett, D. (1991). *Consciousness explained* (pp. 410, 418). New York, NY: Back Bay Books.

[46] Dainton, B. *The phenomenal self* (p. 88). Oxford: Oxford University Press. See also, p. 170; and chapter 6.

[47] Bayne, T. (2010). *The unity of consciousness* (pp. 289–293). Oxford: Oxford University Press.

[48] See, e.g., van Inwagen, P. (2004). The self: The incredulous stare articulated. *Ratio*, *17*(4): 478–491.

[49] Anscombe, E. (1975). The first person. In S. D. Guttenplan (Ed.), *Mind and Language* (pp. 45–65). Oxford: Oxford University Press.

[50] Olson, E. T. (2007). There is no problem of the self. In B. Gertler & L. Shapiro (Eds.), *Arguing about the mind* (p. 262). New York, NY: Routledge.

[51] Lowe, E. J. *Subjects of experience* (p. 32). Cambridge: Cambridge University Press. We are open to Lowe's full thesis of a psychological substance. As it isn't necessary for our arguments, we prefer to limit our commitments.

> My personal identity, therefore, implies the continued existence of that indivisible thing, which I call myself. Whatever this self may be, it is something which thinks, and deliberates, and resolves, and acts, and suffers. I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers. My thoughts, and actions, and feelings change every moment—they have no continued, but a successive existence; but that self or I, to which they belong, is permanent, and has the same relation to all the succeeding thoughts, actions, and feelings, which I call mine.[52]

The self, the subject of consciousness, is a concrete, permanent, naturally occurring substance. The self cannot disappear, and attempts to do so become incoherent. "With the advent of physicalism," writes Chisholm, "the self and its *Beschaffenheiten* [state or condition] seems to be ignored. But it does not take long for the latter to return to the center of the stage."[53]

3.2.3 Subject Necessity

We will use the terms "person," "self," and "subject" interchangeably. We also take the self, the subject, to be the referent of the indexical "I" and its cognates. We will argue for this thesis in Chapter 4. As such, the self can conceive of itself as the unique subject of its thoughts and actions. Contra Hume, we hold, and in Chapter 5 will defend the thesis, that we are directly acquainted with ourselves in states of consciousness.[54] This direct introspectively available knowledge reveals the self as the subject of consciousness. Indeed, a state of consciousness entails a subject of consciousness. "An experience," says Gottlob Frege, "is impossible without an experiencer."[55] More precisely,

> *Subject Necessity*: For all *M*, if *M* is a phenomenal state, then necessarily, there is a subject, *S*, of *M*.

Just as there could not be an instance of motion without a mover, there could not be an instance of a mental property without a mental subject. These commonsense intuitions justify the philosophical intuition that mental properties are kind-defining properties whose instances are substances constituted by those properties.[56] Just as the instantiation of being an electron is a substance constituted by that property as its essence, so the instantiation of mental properties form a spiritual substance constituted by those mental properties. Thus, just as being an electron is an internal constituent exemplified by an electron, so mental properties are internal constituents exemplified by me. And just as the instantiation of being an electron strictly entails an electron, so the instantiation of mental properties strictly entails a mental substance.

---

[52] Reid, T. (1983). Of identity. In R. E. Beanblossom & K. Lehrer (Eds.), *Thomas Reid: Inquiry and essays* (pp. 214–215). Indianapolis, IN: Hackett.

[53] Chisholm, R. (1989). *On metaphysics* (p. 159). Minneapolis, MN: University of Minnesota Press.

[54] Although the direct acquaintance view is popular among dualists, it is embraced by some physicalists or some near enough. See, e.g., Balog, K. (2012). Acquaintance and the mind-body problem. In S. Gozzano & C. S. Hill (Eds.), *New perspectives on type identity: The mental and the physical* (pp. 16–42). Cambridge: Cambridge University Press.

[55] Frege, G. (1967). The thought: A logical inquiry. In P. F. Strawson (Ed.), *Philosophical logic* (p. 27). Oxford: Oxford University Press. First published, 1918.

[56] See, e.g., Lowe, E. J. (2006). *The four-category ontology*. Oxford: Clarendon.

This is the very plausible thesis that all phenomenal states are experiential states of a subject. Consider Galen Strawson's observation.

> Even when all one has assumed to exist is e_e [an individual experience], i.e. that which is given when we consider *e* just in respect of its experiential being, and so without regard to any other being that it may or may not have, one can already know that it involves a subject of experience, simply because experience is essentially experience-*for*.[57]

All phenomenal states, by definition, have a *what-it's-like-to-experience-that-state*; and you only get a *what-it's-like-to-experience-that-state* quality if there is a subject, something that does the experiencing. "The reason for calling it a *phenomenal property*," says Richard Fumerton, "is just that fact—that it is phenomenologically present to *me*."[58] The existence of a phenomenal state entails the existence of a subject. Indeed, it is much more plausible that there are experiences of a subject but without a real-world object.

3.2.4 Subjects, Not Mental States, Are Conscious

We will avoid employing the common conception of a mental state as a state that is itself conscious. We find this misleading, as it disconnects consciousness from subjects of consciousness. In our view, mental states are not themselves conscious. Roderick Chisholm explains,

> The property of being depressed, I have said, is not a property of feeling, and the property of being red is not a property of a sense-datum. But feeling depressed and being appeared to redly are properties of the person—they are nonrelational qualities of the person. So, too, for those other properties which present us with "a qualitative dimension of being.[59]

Agents or subjects are conscious by virtue of being in a mental state. In this view, being conscious is a modification of me, of the subject of experience. I am not conscious derivatively, but primarily, fundamentally. I am not conscious by virtue of some other conscious thing. While this view is uncommon in the philosophy of mind, it does have its defenders.[60]

---

[57] Strawson, G. (2009). *Selves: An essay in revisionary metaphysics* (p. 273). Oxford: Oxford University Press.

[58] Fumerton, R. (2013). *Knowledge, thought, and the case for dualism* (p. 263). Cambridge: Cambridge University Press.

[59] Chisholm, R. M. (1978). Is there a mind-body problem? *Philosophical Exchange*, *9*(1), 28.

[60] See. e.g., Nida-Rümelin, M. (2018). The experience property framework: A misleading paradigm. *Synthese*, *195*, 3361–3387; Nida-Rümelin, M. (2017). Self-awareness. *Review of Philosophy and Psychology*, *8*, 55–82; Peressini, A. F. (2018). There is nothing it is like to see red: Holism and subjective experience. *Synthese*, *195*, 4637–4666; Taylor, H. The relation between subjects and their conscious experiences. *Philosophical Studies*, *177*(11), 3493–3507. For a fascinating argument that the nature of intentionality and mental properties logically require a subject (a substance and not a bundle of properties), see, Chisholm, R. M. (1991). The bearers of psychological properties. *Aaiucov*, *3*, 7–14.

3.2.5 Subject Simplicity and Subject Complexity

Returning to Husserl's mereology, we can divide the logical space regarding the ontology of subjects mereologically. Doing so gives us two theses:

> *Subject complexity*: Subjects of consciousness are mereologically complex—have at least one separable part.

Stated this way, a mereologically complex subject may have both separable and inseparable parts.[61] This is not possible for the following thesis:

> *Subject simplicity*: Subjects of consciousness are mereologically simple—they have no separable parts, are not comprised of separable parts.

According to subject complexity, subjects of consciousness are aggregates with at least one separable part, which stands in external relations.[62] Husserl's distinction between a genuine whole and an aggregate entail that a mereologically simple object is a genuine whole and, therefore, holistically unified. An aggregate, however, a mereologically complex object, is not holistically unified, but is only a unity in thought, it is only functionally unified.

We trust that this ontological detour will help to clarify our discussion in subsequent chapters. The ontology presented here will play a crucial role in the arguments for and against SD, arguments to which we now turn.

---

[61] One reason to think this is important is that, plausibly, anything is an inseparable part of itself.

[62] It is important to note that in some discussions of emergence/complex systems, organized wholes are contrasted with (mere) aggregates. See, e.g., Wimsatt, W. (1995). The ontology of complex systems. *Canadian Journal of Philosophy*, *20*, 564–590.

PART II

ARGUMENTS FROM INTROSPECTION, SELF-AWARENESS, AND INTENTIONALITY

CHAPTER 4

The Real Nature of Introspection Arguments for Substance Dualism

4.1 INTROSPECTION ARGUMENTS FOR SUBSTANCE DUALISM

Introspection provides data — perhaps the most significant data, about the nature of consciousness and its subject. This data, and introspection itself, have proved challenging for physicalism, which some have taken to support arguments for SD.[1] Many SD arguments in the following chapters use introspection and related features of inner-consciousness or self-awareness. The most discussed argument for property dualism, the argument from introspection (AFI), originating in the work of Frank Jackson and Thomas Nagel, faces numerous objections.[2] For this reason, we will spend this chapter addressing one widely popular objection to AFI.

Richard Brandt and Jaegwon Kim made the first contemporary statement of this objection to AFI.[3] Since then, nearly every discussion of AFI has referred to this objection as a decisive refutation of such dualist arguments. Kim evaluates what he takes to be three epistemic arguments for dualism (one from the dubitability of the body, but not the I; one

---

[1] For recent examples see, Fumerton, R. (2013). *Knowledge, thought, and the case for dualism*. Cambridge: Cambridge University Press; and Robinson, H. (2016). *From the knowledge argument to mental substance: Resurrecting the mind*. Cambridge: Cambridge University Press.

[2] See, Jackson, F. (1982). What Mary didn't know. *Journal of Philosophy*, *83*, 291–295; and Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, *83*, 435–450.

[3] Brandt, R., & Kim, J. (1967). The logic of the identity theory. *The Journal of Philosophy*, *64*(17), 515–537.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

from direct access to one's consciousness, but not direct access to a material object or state; and one from private access to one's consciousness, but not to a material object or state), and provides the same solution to each. According to Kim, AFI uses epistemic or intentional properties, such as being known with certainty or being believed to be such and such, to differentiate the mind and body.[4] According to Kim, using epistemic or intentional properties to show that $X \neq Y$ faces problems. In short, from the fact that there are two ways to know something it does not follow that the "items" known are different. Still, Kim admits that "the argument from subjectivity, seems the most compelling, and anyone wishing to reject it should have good reasons."[5]

While we agree that using epistemic properties to show that $X \neq Y$ faces problems, we do not agree that the way Kim's statement of the argument, one that is often employed along with similar points, succeeds. To show this, we will focus on what may be the most developed employment of this line of argumentation to defeat a range of dualist arguments; the treatment of Paul Churchland. As noted earlier, given that several philosophers endorse this line of argument, we take our treatment of Churchland's arguments to apply to all such objections to AFI.

Although Churchland criticizes AFI in various places,[6] we examine his objections in *Matter and Consciousness*, which until now have gone unanswered in the literature. We argue that Churchland's objections are unsuccessful against introspection-based arguments for dualism. Against Churchland's first objection, we argue that it (a) is based on false analogs to AFI, (b) results in introspective skepticism, and (c) presupposes an understanding of AFI that the dualist need not embrace. We offer two alternative accounts of AFI in terms of the linguistic reduction of mental states and the metaphysical relations that mental states stand to their subject. Against Churchland's second objection, we argue that it results in an introspective skepticism that undermines neuroscience, presupposes physicalism, and that Churchland's use of reductions are so unlike mental states that they can't motivate physicalism.

4.2 THE INTENSIONAL FALLACY OBJECTION

Churchland's first objection aims to show that AFI shares the same failures of two arguments he takes as analogous to AFI.[7] He initially states AFI as follows.

AFI_1

1. My mental states are introspectively known by me as states of my conscious life.
2. My brain states are not introspectively known by me as states of my conscious life.
3. Therefore, by Leibniz's Law, my mental states are not identical with my brain states.

---

[4] Kim, J. (2011). *Philosophy of mind* (3rd ed., pp. 153–161). Boulder, CO: Westview Press.

[5] Ibid., p. 161.

[6] Churchland, P. (1985). Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy*, *82*, 8–28; and (2013). Consciousness and the introspection of "qualitative simples." In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience* (pp. 35–56). New York, NY: Springer.

[7] Churchland, P. (2013). *Matter and consciousness* (3rd ed., pp. 50–52). Cambridge, MA: MIT Press.

According to Churchland, AFI_1 commits the intensional fallacy by mistaking different descriptions of or names for the brain (for example, “mind,” “soul,” “spirit”) as picking out something different than the brain when, in fact, they do not.

In making his first objection, Churchland states two arguments he thinks are analogous to AFI_1 yet clearly commits the intensional fallacy. Consider the first.

Ali/Clay:

1. Muhammad Ali is widely known as a heavyweight champion.
2. Cassius Clay is *not* widely known as a heavyweight champion.
3. Therefore, by Leibniz’s Law, Muhammad Ali is not identical with Cassius Clay.

Those conversant with boxing history know that premise (3) is false. “Cassius Clay” is the birth name of the person who changed his name to “Muhammad Ali.” That is, “Muhammad Ali” and “Cassius Clay” are co-referring expressions that pick out the same person. The same issue is at work in Churchland’s second parallel argument.

Aspirin

1. Aspirin is recognized by John to be a pain reliever.
2. Acetylsalicylic acid is *not* recognized by John as a pain reliever.
3. Therefore, by Leibniz’s Law, Aspirin is not identical with Acetylsalicylic acid.

Premise (3) is false. “Aspirin” is the brand name for the same medication to which the technical name “Acetylsalicylic acid” refers. So, again, we have co-referring expressions for the same referent.

Those who do not know that “Ali” and “Clay” or that “Aspirin” and “Acetylsalicylic acid” co-refer to the same person or the same medicine will draw the false conclusion that they are not identical. Churchland claims the same thing is true of AFI_1. Those who do not know that mental states are identical to brain states will be convinced of AFI_1. The force of Churchland’s argument rests on how strongly analogous AFI_1 is to either *Ali/Clay* and/or *Aspirin*. We argue that they are not analogous at all.[8] For the sake of brevity, and because these arguments make the same point, we will refer only to *Ali/Clay*, although our replies apply equally to *Aspirin*.

---

[8] Brian Loar has made a similar objection. Our replies in this section also apply to Loar’s argument. See, Loar, B. (2004). Phenomenal states (revised version). In P. Ludlow, Y. Nagasawa, & D. Stoljar (Eds.), *There’s something about Mary: Essays on phenomenal consciousness and Frank Jackson’s knowledge argument* (pp. 222–223). Cambridge, MA: MIT Press. Using different terminology, Howard Robinson evaluates Churchland’s Ali/Clay objection that we, too, are evaluating. See Robinson, H. (2016). *From the knowledge argument to mental substance* (pp. 17–19). Cambridge: Cambridge University Press. While Robinson grants the value of the argument, he claims that the Knowledge Argument is saved not by spotting an equivocation in Churchland’s response (though he grants there is such an equivocation), but rather, by the totality requirement of the person, e.g., Mary, regarding knowledge of physical facts prior to gaining sight. But we think Robinson misses the point of the arguments below. Our purpose is not to save the Knowledge Argument from introspection—for that task we agree that the epistemic totality of physical facts is crucial. Our aim is more modest. We are seeking to show that Churchland’s counterarguments against Introspective Dualism (which can be expressed in the Knowledge Argument) fail miserably. And we believe Robinson would agree with what follows.

4.2.1 Replies

4.2.1.1 Equivocating Propositional Knowledge and Knowledge by Acquaintance

Ali/Clay equivocates between different types of beliefs (*de re* and *de dicto*) and different types of knowledge (propositional knowledge and knowledge by acquaintance) when compared to AFI. While *de re* belief is belief about an object or agent, *de dicto* belief is belief about the truth of a proposition. Likewise, propositional knowledge is knowledge regarding a proposition, while knowledge by acquaintance is knowledge of some object or agent. These notions of belief and knowledge can come apart. For example, suppose you've never met your neighbor and have no reason to think your neighbor is Buddhist. However, one day at the market you see a person and recognize him as a Buddhist. Moreover, unbeknownst to you, the Buddhist is your neighbor. Consequently, you believe, in the *de re* sense, *of* your neighbor that your neighbor is a Buddhist. However, at the same time you do not believe, in the *de dicto* sense, *that* your neighbor is a Buddhist.

Return to Churchland's parallel arguments. Premise (1) of *Ali/Clay* uses *de dicto* belief. Many individuals believe that Mohammad Ali is a heavyweight champion. Likewise, most people lack the *de dicto* belief that Cassius Clay is a heavyweight champion or that Cassius Clay is Mohammad Ali. Moreover, the type of knowledge in (1) is propositional knowledge, broadly speaking, where S knows that some proposition *p* is true. The same is not true of AFI. On a charitable reading, the dualist understands (1) and (2) in the *de re* sense. Your belief is about your mental state, not about a proposition regarding your mental state. That may come along, but it is not at work in the argument. Likewise, (1) of AFI expresses knowledge by acquaintance, where, broadly speaking, if S has knowledge by acquaintance of *x*, then S is directly aware of *x*.[9]

To be analogous to AFI, *Ali/Clay* must be given the following *de re* reading:

Ali/Clay$_1$

1. People are widely acquainted with that guy [Mohammad Ali] as the heavyweight champion.
2. People are not widely acquainted with that guy [Cassius Clay] as the heavyweight champion.
3. Therefore, Mohammad Ali and Cassius Clay are not identical.

Notice, (2) must be false. Those acquainted with the person they refer to as "Mohammad Ali" are acquainted with the person others refer to as "Cassius Clay." Both names refer to the same person, whom they know. They simply lack the *de dicto* belief that the person they refer to as "Mohammad Ali" is also referred to as "Cassius Clay." However, according to AFI, the same is not true of the mind and brain. Individuals have *de re* beliefs about their mental states by virtue of having knowledge by acquaintance with them. They do not, however, have *de re* beliefs about their brain states, because they lack knowledge by acquaintance with their brain states.[10]

---

[9] In Chapter 5, we will offer a detailed discussion of knowledge by acquaintance with one's consciousness and one's self.

Why think Churchland must present his arguments using *de dicto* belief and propositional knowledge? He must do so because premise (2) of *Ali/Clay* turns out to be false if taken in the *de re* knowledge by acquaintance sense. People who know Mohammad Ali in the *de re* sense also know Cassius Clay in the *de re* sense. Therefore, (1) and (2) must be taken in the *de dicto* sense. Otherwise, they cannot both be true.

Furthermore, notice what happens when *Ali/Clay* is given a *de dicto* reading.

Ali/Clay$_2$

1. "Mohammad Ali" is widely known to refer to the individual who can be described as "the heavyweight champion."
2. "Cassius Clay" is not widely accepted to refer to the individual who can be described as "the heavyweight champion."
3. Therefore, "Mohammad Ali" and "Cassius Clay" are not identical.

Understood this way, *Ali/Clay* is a case of accepting various propositions under different descriptions that refer to the same thing without knowing it. They have certain *de dicto* beliefs but lack others. Moreover, the knowledge in *Ali/Clay*$_2$ is propositional. However, the dualist AFI is in terms of *de re* belief and knowledge by acquaintance, while Churchland's *Ali/Clay* is in terms of *de dicto* belief and propositional knowledge.[11]

Perhaps the physicalist could reply to AFI by rejecting (2). They might do this by maintaining that you are acquainted with your brain states. Notice that this reply does not defend Churchland's objection to AFI. It abandons Churchland's objection and moves to attack premise (2) of AFI. If that is the only response available to the physicalist, then we've successfully shown that the intensional fallacy objection to AFI fails.

4.2.1.2 Absurd Introspection Skepticism

Here is a further reason to think that *Ali/Clay* is not parallel to AFI. The source of error is obvious. We can easily understand how one could fail to know that whom they refer to as "Mohammad Ali" is also referred to as "Cassius Clay." They could simply fail to know his personal history. This error is due to a lack of propositional knowledge regarding the referent of certain names. However, the source of error in AFI is not at all clear. How is it that when attending to and, subsequently, reflecting on my mental states, they seem to stand in relations to me and have certain properties that are not at all like the properties or relations of brain states?

---

[10] We make no general assumption that the only way to have a *de re* belief about something is to have knowledge by acquaintance with it. We claim only that it seems clear that this is what is actually going on in the different cases just discussed. Of course, Churchland would not agree with our assessment because he explicitly denies that there is any such thing as simple seeing or knowledge by acquaintance. For Churchland, all seeing is theory-laden and amounts to seeing as or seeing that (See, Churchland, *Matter and Consciousness*, 130). We ask the reader to look carefully at the cases, and judge who has more accurately captured the dialectic.

[11] Although Churchland does accept knowledge by acquaintance, he has posited a neuroscientific distributive representational theory of acquaintance. See Churchland, P. (1989). Knowing qualia: A reply to Jackson. In P. Churchland (Ed.), *A neurocomputational perspective* (p. 71). Cambridge, MA: MIT Press.

Someone could object that mental states seem to have these non-physical features only if one has already accepted dualism, while a physicalist will not see those states as having these features. However, as Kim has correctly noted, this response has it backward. In context, Kim is talking about various dualist intuitions expressed in a host of thought experiments that seek to provide evidence that qualia are non-physical and contingently related to physical states. Here is how he concludes his discussion of these issues:

> The [property-dualist] case against qualia supervenience therefore is not conclusive, though it is quite substantial. Are there, then, considerations in favor of qualia supervenience? It would seem that the only positive considerations are broad metaphysical ones that might very well be accused of begging the question.[12]

The only account Churchland could give assumes that we cannot know the basic nature of our own mental states by paying attention to them. But this kind of skepticism about introspection is absurd. If we cannot know the basic nature of our own mental states by introspection, then we cannot know which thoughts are our thoughts. As John Haldane writes, "How am I even to pose the question, let alone devise and operate procedures for determining whether it was me or someone else who was thinking unless I already have a grasp of certain thoughts as being mine?"[13]

Likewise, how can you know that you are in a state of pain or a state of joy if you cannot attend to your mental states by reliable introspection? Imagine you are convinced that pain is identical to a specific brain state. While hooked up to and observing your own fMRI scan, you bite your tongue. At the same time, you notice that your brain states are not behaving in the way associated with pain. It would be truly odd for you to exclaim, "I am having an experience of pain, but I am not in pain."

One might object that although we know something about the nature of our mental states, there is much we don't know by introspection. For instance, you cannot tell by introspection whether mental states are fundamentally your states, or whether they are derivatively your states but grounded in the states of something other than you. The temporality of mental states is obscure. How long does a state last? The composition is obscure. When I hear "Paris," is that a holistic experience or do I experience five sounds? No one is surprised that you cannot know the nature of the sun just by paying attention to it. Why should our mental states be different?

We grant that we don't know *everything* about our mental states. However, we clearly know certain things about them that are relevant to our argument against Churchland, viz., their intrinsic phenomenal content (the what-it-is-like-to-be-in-pain, that the content of my thought is "that London is beautiful" and not "that I hate cabbage," that they belong to me). Our argument does not assume that we know everything about our mental states; it merely presupposes we know things about their intrinsic nature such as the items just listed. That is sufficient for our argument to go through.

Churchland could reply that all we have shown is that one cannot know the physical nature of a mental state by introspection, but it still has a physical nature. We offer three responses. First, as Richard Swinburne points out, at least certain mental states (e.g., thoughts) are conscious events of which I am aware. Moreover, “there is no more to a subject’s thoughts than they are aware of.”[14] They are, as it were, “naked before the subject.” Second, it is very difficult to take most, if not all, mental states as having a physical nature. Thoughts, beliefs, desires, and exercises of volitional power do not seem to have geometrical properties (location, size, shape), mass, charge, and so on. Indeed, it seems to be a category fallacy to ascribe such properties to these mental states. It would seem impossible to find such features because any candidate for the physical nature of a mental state could easily be treated as a physical causal factor that must be present for the mental state to obtain. Finally, even if we grant, for the sake of argument, that mental states have a physical nature, our arguments still show that there is an irreducible mental nature to them as well. This is all we need for our arguments to succeed.

---

[12] Kim, J. (2006). *Philosophy of mind* (2nd ed., p. 233). Boulder, CO: Westview Press.

[13] Haldane, J. (2003). "(I am) thinking." *Ratio*, *XVI 2*(June), 135.

Furthermore, such examples rest on an equivocation regarding “whose thought something was.” The phrase can be causal or authorial. In this sense, a thought is one’s own just in case that person caused or authored the thought. In this case, confusion seems quite plausible. For example, one could often fail to tell if a thought came from or was authored by their twin rather than coming from one’s self. The phrase can also mean, “belongs to or had.” In this sense, thought is one’s own, just in case it is possessed by or had by that person. More precisely, a thought is one’s own just in case the mental property, which consists of the immanent essence/propositional content of the thought, is exemplified by that person to form a mode of the person: the-having-of-the-mental-property-by-Smith. It is far from clear that a person could be confused about a thought belonging to herself or someone else. These counterexamples gain their force by being examples of the first sense and not the second sense of the phrase.

Lastly, one might reply that we cannot know the full nature of a mental state by introspection. This leaves open the possibility that we can typically know enough to tell that it is a state of pain, even though we cannot know through introspection that it is a brain state. But this reply won’t work. We have already looked at non-question-begging reasons to support the view that what we do know about our mental states is their intrinsic nature, which helps to ground their cross-world identity. And this nature is just not physical. Things such as *what-it-is-like*, being subjective, first-person, perspectival, being of, having semantic content, and so on play absolutely no role in physical theory. One will search in vain through a physics, chemistry, or neurophysiology text in an attempt to find these notions listed among the properties that are counted as physical.

Of course, there are things we don’t know about our mental states through introspection. It is hard to tell precisely how long a mental state lasts or all the entailments that follow from it. However, it could be argued that both problems follow from the mental nature of consciousness and its states. Unlike sequences of physical states, consciousness moves in a continuous flow, so it is hard to draw sharp temporal boundaries to a mental state. But the movement of brain states consists of a series of discrete states that have clear boundaries and that jump from one to the other. Regarding entailment, arguably, physical states stand in causal relations to each other, but not in intentional, epistemic, or logical relations. But these are precisely what characterize the relations among the immanent propositional content of propositional attitudes.

---

[14] Swinburne, R. (2013). *Mind, brain, and free will* (p. 79). Oxford: Oxford University Press.

But in the epistemic sense, couldn't it be the case that the unknown aspects of our mental states include physicality? In our view, the only positive argument for this claim is a question-begging prior commitment to physicalism. There is no other positive evidence for it.

4.2.1.3 Missing the Target: Why AFI Is Metaphysical

Perhaps the greatest problem is that Churchland has failed to represent AFI accurately. Churchland's objection assumes the dualist is committed to the following understanding of AFI.

> $_{Epistemic}$ AFI: Two different ways of knowing something shows the things known are different.

From this assumption Churchland charges the dualist with deriving a metaphysical conclusion from epistemological premises. But why think that is the case? Why couldn't the dualist be committed to the following understandings of AFI?

> $_{Language}$ AFI: When I introspect, the natures of the items given to me are not describable in physical language.

This is at least one way Howard Robinson understands AFI. In one statement of the knowledge argument, Robinson concludes that "The phenomenal nature of sound cannot in principle be expressed in the vocabulary of physical science."[15] This is quite different from Churchland's presumption of $_{Epistemic}$ AFI.

One might reply that the nature of a statue, for example, is not describable in physical language (it requires the language of art criticism). But that does not mean that the statue isn't a physical thing. We think this objection rests on a mistake about what makes something a statue. A statue is a certain amount of stuff, say clay, shaped in a certain way and intended by its maker and/or the community of people who intend to take the shaped clay to represent what the maker intended. It is the maker's intentions, along with those of her community, which makes the object a statue. And, arguably, intentions are mental entities. Consider a possible world with no sentient beings. There would be no statues in this world, just shaped stuff. If erosion created an object shaped like the Statue of Liberty, it would merely be shaped stuff and not a statue.

In addition to $_{Language}$ AFI, there is a further dualist understanding of AFI:

> $_{Metaphysical}$ AFI: Things that can be known by introspection (mental states and properties), given what introspection is, stand to the subject of experience as (a) inner (not spatially inside, but properties they exemplify); (b) private (known by the subject in a way not available to anyone else); (c) directly given to the subject and not by means of anything else (self-presenting); (d) intentional; and (e) necessarily, states of the subject. None of these features is true of all or at least some brain states.

---

[15] Robinson, H. (1993). The anti-materialist strategy and the "knowledge argument." In H. Robinson (Ed.), *Objections to physicalism* (p. 163). Oxford: Clarendon.

John Searle expresses a similar view:

> The logical form of the argument is this: I stand in a relation to certain entities, my experience of colors. And the bat stands in a relation to certain entities, its experiences of what it feels like to be a bat. A complete third-person description of the world leaves out these entities, therefore the description is incomplete.[16]

Richard Fumerton's understanding of the knowledge argument also accords well with $_{\textit{Metaphysical}}$ AFI.[17] $_{\textit{Metaphysical}}$ AFI focuses on unique metaphysical features of mental states. Mental states are the sorts of things that can be introspectively known; thus, they stand in a direct access relation to the subject.[18] Mental states are inner—exemplified within the subject—and are necessarily states—modes—of the subject.[19] Mental states are intentional; they are about something other than themselves. Brain states, however, lack these features.[20] We have, then, a second understanding of AFI that eludes Churchland's first objection.

If Churchland is to justify this presumption, we see only two moves available to him. First, he could argue that dualists use $_{\textit{Epistemic}}$ AFI. Churchland fails to do so, and we are unaware of anyone who defends $_{\textit{Epistemic}}$ AFI. However, suppose Churchland pointed to dualists who do. What follows? At best he has retired a bad version of AFI. That would be helpful if such arguments are being made. As we have stated this is unlikely. Furthermore, this move fails to offer an objection to AFI in general. Given that Churchland's aim is to do just that, this response is unhelpful.

Churchland could argue that $_{\textit{Language}}$ AFI and $_{\textit{Metaphysical}}$ AFI are reducible to $_{\textit{Epistemic}}$ AFI. This would be a much stronger argument. However, this line of argument hardly seems possible. The conclusions of $_{\textit{Language}}$ AFI differ from both $_{\textit{Metaphysical}}$ AFI and $_{\textit{Epistemic}}$ AFI. Moreover, the conclusions of each argument could not be derived from their premises. This line of reasoning is unpromising.

4.3 THE OBJECTION FROM NEUROLOGICAL *DE DICTO* BUT NOT *DE RE* BELIEFS

To make his second objection, Churchland offers another argument he claims is analogous to AFI.

AFI$_2$

1. My mental states are knowable by introspection.
2. My brain states are not knowable by introspection.
3. Therefore, by Leibniz's Law, my mental states are not identical with my brain states.

---

[16] Searle, J. R. (2004). *Mind: A brief introduction* (p. 68). Oxford: Oxford University Press.

[17] Fumerton, R. *Knowledge, thought, and the case for dualism*, chapter 3.

[18] See, Moreland, J. P. (2003). The knowledge argument revisited. *International Philosophical Quarterly*, *43*(2)(170), 219–238.

[19] See, Lowe, E. J. (1996). *Subjects of experiences*. New York, NY: Cambridge University Press.

[20] See, Bealer, G. (1993). Materialism and the logical structure of intentionality. In H. Robinson (Ed.), *Objections to physicalism* (pp. 101–126). Oxford: Clarendon; and Lowe, E. J. (1995). There are no easy problems of consciousness. *Journal of Consciousness Studies*, *2*(3), 206–271.

While AFI_1 made introspectability a feature of the way something is known, AFI_2 makes it a property of the mental state itself. Churchland's strategy is to reject AFI_2 by denying premise (2). He explains,

> For if mental states are indeed brain states, then it is really brain states that we have been introspecting all along, though without fully appreciating what they are. And if we can learn to think of and recognize those states under their familiar mentalistic descriptions, then we can certainly learn to think of and recognize them under their more penetrating neurophysiological descriptions. At the very least, premise (2) begs the question against the identity theorist.[21]

The claim here is that we have *de re* beliefs of our brain states and lack *de dicto* beliefs that what we are aware of in introspection are brain states. Consequently, we do have knowledge by acquaintance with brain states. We simply lacked the propositional knowledge that they are, in fact, brain states.

To illustrate this second objection, Churchland offers the following argument, which he takes to be parallel to AFI_2:

Mean KE
1. Temperature is knowable by feeling.
2. Mean molecular kinetic energy is not knowable by feeling.
3. Therefore, by Leibniz's Law, temperature is not identical with mean molecular kinetic energy.

Churchland holds that this argument is unsound by denying premise (2). He explains:

The culprit, as above, is premise (2). Just as one can learn to feel that summer air is about 70 °F, or 21 °C, so one can learn to feel that the mean KE of its molecules is about 6.2×10^{-21} joules, for whether we realize it or not, that is what our native discriminatory mechanisms are keyed to. Perhaps our brain states are similarly accessible, if we learn the relevant theoretical vocabulary, and if we learn to spontaneously apply the relevant terms.[22]

Churchland's rejection of (2) is that it is possible that, given certain scientific advances, we come to know by feeling that mental states just are brain states, that the mental state of feeling the summer air is identical to some brain state. We have three responses to this objection.

4.3.1 Replies

4.3.1.1 The Neuroscientific Loss of Introspective Skepticism

Mean KE comes at a great cost. If Churchland is correct, it becomes utterly mysterious how you could know your own mental states by introspection since for Churchland they are completely different from what they seem to be in introspection. In fact, the radically disparate nature of what is presented to us in introspection regarding the nature of consciousness is at the heart of entrenched disputes in philosophy of mind (e.g., the explanatory gap argument). We argued above (4.2.1.2) that this leads to the problem of not being able to know what thoughts are your thoughts. One could respond that a non-deviant causal chain between a scanning region of the brain and the brain state being scanned solves this mystery. We provide a critical evaluation of this move in Chapter 10. Here we raise two further problems.

---

[21] Churchland, pp. 51–52.

[22] Churchland, p. 52.

According to many dualists, nothing we know by introspection can be described in reductive physical language. The argument is that what we know by introspection is constituted by qualia. Qualia, the dualist contends, resist even linguistic reduction. This is the argument made in $_{\textit{Language}}$ AFI. It seems then Churchland's objection will be of little use against these dualists. It merely gets us back to the issue at hand.

Furthermore, Churchland's introspective skepticism undermines the empirical research of many neuroscientific studies. For example, a 2004 study discovered that regret correlates with the brain's orbitofrontal cortex.[23] Subjects were given a choice between two risky gambles. In partial feedback blocks, the result of the gamble appeared only in the gamble they selected. In complete feedback blocks, the results of both gambles appeared. Subjects reported a range of sadness or disappointment when they lost the bet. Those who also discovered that they could have won, had they chosen the other betting option, reported a degree of sadness or regret. Researchers discovered that subjects with lesions to the orbitofrontal cortex did not experience sadness after learning they could have won had they chosen differently, hence, no experience of regret. In these experiments researchers must ultimately rely on the introspective first-person report of the subject regarding their mental states, which allows us to distinguish between regret and disappointment.

One might reply that in asserting that Churchland's view of mental states is completely different from what they seem to be in introspection, we are begging the question. A materialist might hold that Churchland's account of mental states is not different, just richer. For example, there is so much more to a table than meets the eye (all that stuff going on at the molecular and atomic level).

Unfortunately, this is not an option for Churchland. For one thing, Churchland is an eliminative materialist regarding propositional attitudes. He recommends we search for brain states that, as long as they play the "same role" as propositional attitudes as traditionally conceived, they need bear no resemblance whatsoever to those propositional attitudes. Moreover, he simply identifies qualia with brain states. It is not as though a quale is irreducibly mental but there is a vast unknown aspect to it that is physical. No, if qualia are allowed to exist as irreducible entities, they are so different from brain states that the physicalist will have a difficult time with the location problem. It is in no way question-begging to assert that the phenomenal texture that constitutes a quale is utterly unique compared to anything else in the universe and phenomenal texture is just not among the properties, relations, or parts that constitute physical science.

Indeed, it is the physicalist who begs the question here. In the context of debates about property dualism regarding qualia, Kim correctly notes the following.

> The [property-dualist] case against qualia supervenience therefore is not conclusive, though it is quite substantial. Are there, then, considerations in favor of qualia supervenience? It would seem that the only positive considerations are broad metaphysical ones that might very well be accused of begging the question.[24]

---

[23] Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J.-R., & Sirigu, A. (2004). The enjoyment of the orbitofrontal cortex in the experience of regret. *Science*, *304*, 1167–1170.

[24] Kim. *Philosophy of mind* (2nd ed., p. 233).

4.3.1.2 The Illicit Presumption of Physicalism

Churchland's rejection of (2) assumes that mental states are identical to brain states. If they are not, then it is impossible to come to know by introspection that mental states are just brain states. However, this is the proposition in dispute. We see only two plausible ways for Churchland to deny (2). First, argue that brain states are identical to the kinds of things we know by introspection. Second, argue that brain states share the relevant features that make them knowable by introspection. The most plausible candidates in this discussion are mental states and their features, such as being self-presenting, which make them knowable by introspection. Yet, Churchland's rejection of (2) merely assumes this. That is, Churchland's objection to AFI_2 in terms of *Mean KE* begs the question against the dualist.

4.3.1.3 Dubious Analogies from the History of Science

Perhaps one could reply that Churchland is not presuming physicalism in *Mean KE*, but rather importing the promises of science to one day reduce mental states to brain states. Churchland certainly uses this line of reasoning in his overall case for reductive physicalism. He submits the following list of reductions in the history of sciences.

(i) Sound is just a train of compression waves traveling through the air.
(ii) Light is just electromagnetic waves.
(iii) Color (of an object) is just a triplet of reflectance efficiencies (of the object).
(iv) Warmth is identical with low average molecular kinetic energy.

Churchland surmises that the success of science in discovering these reductions justifies the supposition that science will someday show that mental states are identical with brain states. Consequently, (2) of *Mean KE* will one day be falsified.

We offer two ways of responding. First, one could challenge the legitimacy of the examples of ontological reduction. Some, such as Steve Horst, have argued that the alleged historical parallels of reduction and identity are questionable. Horst points out that contemporary philosophy of mind is

> predicated upon the assumption that intertheoretic reductions are widespread at the junctions between the various natural science and are available in principle at all such junctions … Curiously, though, at the time that reductionism was reaching its zenith in philosophy of mind, these assumptions were coming under fire within philosophy of science itself. Today, these assumptions are in fact generally rejected by philosophers of science.[25]

We are only suggesting that this line of reply seems plausible. It is beyond the scope of this chapter to defend Horst's thesis. Although Horst is not alone in his view, some will not be persuaded by this reply. The following may persuade a larger audience.

Suppose that Churchland's cases of reduction are accurate. One could still reply that they are bad analogies for mental properties. Consider heat and color. One could grant that there is a difference between what they are (mean kinetic energy and wavelengths) and how they appear (as warm, as red). That is, heat and color pose both phenomenal properties and physical properties. However, such a distinction is not available for mental states like a sensation of pain. A pain sensation is just the way it appears to a subject, viz., as hurtful. The nature of a painful feeling and the appearance of pain do not come apart.[26]

---

[25] Horst, S. (2007). *Beyond reduction: Philosophy of mind and post-reductionist philosophy of science* (p. 47). Oxford: Oxford University Press.

Imagine that an individual's brain is in a neural state normally correlated with a sensation of pain. It seems possible that they could be in such a state and not have a sensation of pain. However, it does not seem possible for an individual to have a sensation of pain and not experience it. Take away the experience of pain and pain goes with it. This is at the heart of why nearly all philosophers have rejected behaviorism. Hilary Putnam's "X-worlders" or "super-super-spartans" can simultaneously experience pain and suppress all pain behavior by sheer force of will.[27] Likewise, we can imagine, what Ned Block calls, "perfect pain-pretender" who, although they are experiencing no pain at all, act exactly as one who is in severe pain.[28] These thought experiments show that the phenomenal texture of a pain sensation is essential to what a pain sensation is.

Consequently, mental states, like pain, resist the type of reduction to which things like sound, light, and heat are vulnerable. One may concede that the history of science includes several successful reductions and, at the same time, reject the plausibility of reducing mental states to physical states.

4.3.1.4 Grahek's Challenge to the Phenomenal Essence of Pain

We have been assuming the commonsense subjectivist view of pain according to which the essence of a pain is the phenomenal, hurtful what-it-is-like to a pain, along with pain's subjective nature (its first-person, subjective, perspectival aspect). However, following an important 2001 book published by the late Nikola Grahek, some philosophers believe there is hard scientific evidence against the subjectivist view.[29]

To understand Grahek's argument, we first must get clear on what he calls the two components or dimensions to pain. Grahek correctly notes that pain is a complex experience with different components. First, there is "feeling pain," which refers to the sensory-discriminatory aspects of pain. These are the qualitative hurtful qualitative aspect to pain (the what-it-is-like to have a sensation of pain) that can be experienced to varying degrees of intensity. Second, there is "being in pain" which refers to the affective-motivational aspect of a pain perception, basically the unpleasantness of pain.

According to Grahek, one can feel pain without being in pain and the only clear example of this is people with Pain Asymbolia/Disassociation. Due to various physiological defects, such people acknowledge feeling pain but it is not unpleasant to them. They lack an experience of the affective-motivational aspect of pain. On the other hand, one can be in pain without feeling pain. In Chapter 7, Grahek reports on a patient with various brain lesions whose left hand was exposed to a laser that usually brings about a spatially precise feeling of heat pain. The patient described an intensity-sensitive, unpleasant feeling somewhere between his fingertips and his shoulder that he desired to avoid. He could not provide further qualitative descriptions of the feeling (e.g., being warm, hot, mildly painful, intensely painful). This patient had the affective-motivational aspect of pain but not the sensory feeling of pain.

---

[26] As Saul Kripke points out, heat, for example, is picked out by a reference-fixing description that it only contingently satisfies ("that which causes heat sensations"), which accounts for the illusory "appearance of possibility" that there could be heat without molecular motion. But we don't pick out pain by way of any contingently associated property of pain.

[27] Putnam, H. (1975/1963). Brains and behavior. In *Mind, language, and reality: Philosophical papers,* (Vol. 2, pp. 325–341). Cambridge: Cambridge University Press.

[28] Block, N. (1981). Psychologism and behaviorism. *The Philosophical Review*, *90*, 5–43.

[29] Grahek, N. (2007). *Feeling pain and being in pain* (2nd ed.). Cambridge, MA: MIT Press.

It seems to us that these cases do nothing to show that the subjective view of pain we employ is false or problematic in some other way. Beginning with the patient just described, all the case shows is that the affective-motivational component that usually accompanies pain sensations is not sufficient for sensing pain. Why? Because the patient lacks something essential to pain, viz., the phenomenal, sensory-discriminative feeling that is essential to a mental state being a pain.

Grahek's use of people with Pain Asymbolia to refute the subjective view of pain is a failure. In these cases, Grahek acknowledges that people experience the sensory-discriminatory component of pain. But with the second, unpleasantness component gone, the "pain" state does not satisfy the normal functional role of pain. After all, the subjectivist view of pain eschews functional, relational accounts of pain. So, a failure to satisfy the normal functional role of pain is of little consequence. Given that we agree that pain is a complex experience, it is consistent with Grahek's data that while the sensory-discriminatory component of pain is essential, the affective-motivational aspect is accidental. However, in the vast majority of cases, it is present. We conclude that Grahek-style arguments fail to undermine the subjective view of pain.

4.4 SEARLE'S ANTI-INTROSPECTION ARGUMENT

John Searle has raised the following argument against introspection and self-awareness that needs to be addressed:[30]

> (1_I) If the standard model (of awareness) is true, then there is a distinction (presumably, not a distinction of reason) between the thing seen and the seeing of it.
> (2_I) The standard model is true.
> (3_I) Therefore, there is a distinction between the thing seen and the seeing of it.
> (4_I) If introspection occurs, then there is no distinction between the thing seen and the seeing of it.
> (5_I) Therefore, introspection does not occur.

There are at least two problems with this argument. Let us begin with (4_I). Searle gives no good reason to accept (4_I) and there are good reasons to reject it. Let us assume, as is usually granted, that in introspection we often have a second-order mental state directed upon a first-order mental state. For example, in introspection, the self—whatever it is—is directly aware of a sensation of red or a feeling of pain. This is a perfectly intelligible account of introspection, and it provides the distinction required to reject (4_I).

---

[30] Searle, J. (1992). *The rediscovery of the mind* (pp. 96–97, 143–144, 149). Cambridge, MA: MIT Press.

If one rejects this model of introspection, one can still rebut Searle's argument by rejecting (2_I). That is, one can grant (1_I) for the sake of argument and deny that it applies to introspection on the grounds that it begs the question to assert such an application. After all, why apply the standard model to introspective acts? Certain mental states relevant to introspection are self-presenting properties. And according to a standard characterization, a *self-presenting property* presents to a subject, the intentional object of that property (for example, an apple's surface), and the self-presenting property itself (being-an-appearing-of-red). Such properties present other things to a subject intermediately by means of them, and they present themselves to a subject directly simply in virtue of the fact that he has them. Introspective awareness of being-an-appearing-of-red could be understood as the exemplification of a self-presenting property.

In this case, introspection provides a counterexample to the standard model. And while Searle does not mention the self, we see no reason why one cannot be directly aware of oneself. In Chapter 5 we will argue that we are directly self-aware. An aspect of that argument employs a Husserlian understanding of intentionality, not as a relation, but as a *monadic property*. Accordingly, when one is aware of oneself (as opposed to a mental state one has), in direct self-awareness, one simply directs one's intentionality onto oneself and the subject and object of awareness stand in the identity relation to each other. We mention this briefly to show that nothing Searle says comes close to undermining such an understanding of self-awareness.[31]

Searle similarly attacks a spatial metaphor associated with "privileged access" that he alleges to go proxy for introspection:[32] When I spatially enter something, there is a distinction among me, the act of entering, and the thing entered. No such distinction obtains in alleged acts of "private access"; thus, "private access" should be rejected. The appropriate rebuttal follows the lines of response given to the argument against introspection.

4.5 CONCLUSION

The goal of this chapter was to vindicate dualist arguments from introspection against a popular objection defended by most physicalists. To that end we have offered several arguments. Against the first objection, we argued that it is based on two dissimilar arguments, as they employ different notions of belief and knowledge than AFI. Second, it results in an absurd introspective skepticism. Third, the objection presupposes an understanding of AFI that the dualist needn't embrace. Against the second objection we offered three replies. First, it results in a costly introspective skepticism that undermines reliance on neuroscience. Second, this objection presupposes physicalism. Third, the supposed reductions in the history of science may not turn out to be reductions, and if they are they are so unlike mental states that they give us no reason to expect that mental states are identical to brain states. Consequently, such objections needn't seduce dualists and physicalists, and their popularity in the literature should fade.

---

[31] For an insightful criticism of Searle's theory of intentionality, see, Meixner, U. (2006). Classical intentionality. *Erkenntnis*, *65*, 25–45.

[32] Searle, p. 98.

CHAPTER 5

From Self-Awareness and Intentionality to the Self as Soul

Here are two facts widely acknowledged even by steadfast SD detractors: physicalism seems false to nearly everyone, while substance dualism seems obviously true.[1] Thus, Jaegwon Kim's concession:

> We commonly think that we, as persons, have a mental and bodily dimension ... Something like this dualism of personhood, I believe, is common lore shared across most cultures and religious traditions ...[2]

Unlike physicalism, we needn't be trained to be dualists. The data below shows that dualism, especially SD, is a core, base, commonsense belief resulting from phenomenologically thick intuitions or seemings. That we are "bodies inhabited and governed in some intimate if mysterious way by minds (or souls)," says William Lyons, "seemed and still seems to be nothing more than good common sense."[3]

In this chapter we focus on the phenomenologically and empirically established fact that dualism in general and SD in particular seem true, and what we will call *dualist seemings* are ubiquitous, diverse, and hard to deny. The best explanation of this large data

---

[1] Among these philosophers, Charles Taliaferro includes Michael Levin, Daniel Dennett, David Lewis, Thomas Nagel, J. J. C. Smart, Richard Rorty, Donald Davidson, and Colin McGinn. See Taliaferro, C. (2001). Emergentism and consciousness. In K. Corcoran (Ed.), *Soul, body and survival* (p. 60). Ithaca, NY: Cornell University Press.

[2] Kim, J. (2001). Lonely souls: Causality and substance dualism. In K. Corcoran (Ed.), *Soul, body and survival: Essays on the metaphysics of human persons* (p. 30). Ithaca, NY: Cornell University Press.

[3] Lyons, W. (2001). *Matters of the mind* (p. 9). New York, NY: Routledge.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

set is that we are directly aware of our self as a mereologically simple, immaterial substance. We begin with an intuition pump developed by David Barnett and argue that the extant objections to this argument each fail. The bulk of the chapter explains why certain kinds of intuitions—what we will call dualist seemings—are a source of data that cannot be ignored and is best explained by SD.

5.1 DAVID BARNETT'S SIMPLICITY ARGUMENT

In what follows, we analyze David Barnett's argument for the plausibility of the thesis that subjects of consciousness must be mereologically simple. Barnett argues that what we call a dualist seeming (*xi*) *Simple Self* (that pairs of subjects of consciousness cannot themselves be a subject of consciousness) is plausibly explained by the fact that subjects of consciousness must be mereologically simple. We aim to briefly present Barnett's argument, analyze the extant replies, and argue that these alternative explanations fail.

5.1.1 Barnett's Argument Stated

David Barnett has argued in several places that subjects of consciousness cannot be composed of proper parts (plausibly what we call separable parts) but must be mereologically simple, from what he takes to be an uncontroversial claim:

> *The Datum*: For any pair of conscious beings, it is impossible for the pair *itself* to be conscious.

Imagine that we form a person-pair. According to *the Datum*, it is impossible that the person-pair we form is a subject of phenomenal unity. Barnett distinguishes six possible explanations of the *Datum*, which are intended to cover the logical space:

Number: Person-pairs lack a sufficient *number* of immediate parts.
Relation: Person-pairs lack immediate parts capable of standing in the right *relations* to each other.
Nature: Person-pairs lack immediate parts of the right *nature*.
Structure: Person-pairs are not proper *structures*.
Combination: A combination of Number, Relation, Nature, and Structure.
Simplicity: Pairs of people are not simple.

The first five accounts, argues Barnett, do not explain the *Datum*, whereas *Simplicity* does. Formally stated, Barnett's argument is as follows.[4]

(S1): The *Datum*: for any pair of people, it is impossible for the pair itself to be conscious.
(S2): Something must explain the *Datum*; there is at least one non-trivial feature that no pair of people could itself have but which every conscious being must have.
(S3): If no combination of *Number*, *Relation*, *Nature*, and *Structure* explains the *Datum*, then Simplicity explains the *Datum*.

---

[4] Barnett, *You are simple*, p. 172.

(S4): No combination of *Number*, *Nature*, *Relation*, and *Structure* explains the *Datum*.
(S5): Therefore, *Simplicity* explains the *Datum*.

According to Barnett, (S1) and (S2) are obviously true, while (S3) is justified by the failure of the alternative explanations of the *Datum*. Premise (S4) is, therefore, the contentious premise.

Barnett's defense of (S4) pumps the intuition that *Simplicity* is at least plausible. For our purposes we needn't consider how Bailey executes his project. Our purpose is to consider rival explanations of the *Datum* that are presented as missing from Barnett's rival explanations of *Simplicity*.

5.1.2 Objections to Barnett's Argument

Few have raised objections to Barnett's argument. Those who have argue that Barnett has overlooked alternative explanations of the *Datum*. In this section, we will explain and reply to each.

5.1.2.1 Elimination

Andrew Bailey argues that the *Datum* can be explained just as well by the following thesis that not only rivals but is superior to *Simplicity* as it is more plausible given materialism as material simples seem unlikely. *Elimination*: Person-pairs are metaphysically impossible. Bailey motivates *Elimination* by assuming that a person-pair must be a fusion of persons. Of course, if person-pairs cannot because the fusion of persons is impossible, this would explain the Datum.

Bailey invokes the standard notion of mereological fusion.[5]

The classical mereology (Stanislaw Lesniewski, Nelson Goodman, W. V. Quine, and David Lewis) employs, among others, the following three mereological notions.[6]

Part: x is a part of $y =_{\text{def.}} x$ is a proper part of y and $x \neq y$.
Overlap: x and y overlap $=_{\text{def.}}$ for some z, z is a part of x and a part of y.
Compose: The fs compose x (x is a fusion of the fs) $=_{\text{def.}}$ Every f is a part of x, and every part of x overlaps an f.

Together these entail that if x is a fusion of the fs then every f is a part of x, and every part of x (the fs) overlaps an f. We add the following boundary condition.

Boundary: if x overlaps an f, then there is no boundary between x and f.

Taken together, these four mereological theses entail:

Fusion: x is a fusion of the fs $=_{\text{def.}}$ every f is a part of x, and every part of x (the fs) overlaps an f, such that there is no boundary between x and f.

---

[5] Bailey cites van Inwagen, P. Can mereological sums change their parts? *Journal of Philosophy, 103*(12), 614–630.

[6] Gideon Rosen and Cian Dorr, "Composition as a fiction," in Richard M. Gale (Ed.), *The Blackwell guide to metaphysics*. Oxford: Blackwell, 2002, 152–153.

Applied to person-pairs, we get:

> *Person-Pair Fusion*: A person-pair (PP) is a fusion of person$_1$ and person$_2$ (P_1 and P_2) $=_{\text{def.}}$ P_1 and P_2 are a part of PP, and every part of PP overlaps P_1 and P_2, such that there is no boundary between P_1 and P_2.

According to Bailey, while there are principles governing when there is a fusion of some objects, "No correct principles of this kind—no correct answer to the so-called special composition question—yields the result that there are or could be person-pairs."[7] This, says Bailey, explains the *Datum*.

But why suppose, as *Elimination* does, that *x* is a person-pair, *iff*, *x* is a fusion of two persons? And why suppose that each person of a person-pair is an immediate proper part of the fusion? As we understand it, mereological fusion is not a case of two things becoming identical with each other or identical with some other thing. Rather, in a mereological fusion things become parts of the fused thing.[8] Therefore, two persons become proper parts of a whole they compose. Following Roderick Chisholm, we understand person-pairs as heaps or aggregates.[9] It seems the members of a fusion are *Separable Parts* as they can exist and retain their identity outside the fusion.[10] A mereological fusion cannot be a *Genuine Whole*. Therefore, person-pairs do not form a holistic unity but only a functional unity.[11] Moreover, if we reject eliminativism for heaps and aggregates by granting that in some sense or other, they are wholes in some sense, they are still not the right kind of wholes to be subjects of consciousness having their own totalizing what-it-is like. Why? Because they are not simple, just as Barnett concludes. Finally, this means that *Elimination*, rather than being an overlooked explanation of the *Datum*, plausibly turns out to be a version of *Relation*.

5.1.2.2 Maximality

Bailey's second account of the *Datum* is that consciousness is a maximal property. The notion of maximality is important, and there seem to be two different, closely related maximality principles relevant to Bailey's objection.

> *Individual Maximality* (IM): *x* is an individual maximal object *iff* necessarily, *x* is an F and there are no proper parts *p* of *x* that are Fs.[12]

> *Property Maximality* (PM): A property P is maximal *iff* necessarily, for any *x* that is F, there is no *y* such that *x* is a proper part of *y* and *y* is F.

---

[7] Ibid.

[8] Chisholm, R. M. (1979). *Person and object: A metaphysics study* (p. 93). Chicago, IL: Open Court.

[9] Chisholm, R. M. (1991). On the simplicity of the soul. *Philosophical Perspectives*, *5*, 172. Heaps, aggregates, and homeodynamic systems will be analyzed in Chapter 8.

[10] *Separable part*: *x* is a separable part of some whole *W iff* (i) *x* is a particular, (ii) *x* is a part of *W*, and (iii) it is possible for *x* to exist without being a part of *W*. *Inseparable part*: *x* is an inseparable part of some whole *W iff* (i) *x* is a particular, (ii) *x* is a part of *W*, and (iii) it is not possible for *x* to exist without being a part of *W*. See Chapter 3.

[11] *Functional Unity*: O is functionally unified $=_{\text{def.}}$ (i) O is an aggregate, and (ii) the separable parts of O work together to cause some effect as a consequence of the external relations among O's parts. See Chapter 3.

[12] *Individual Maximality* is embraced, for example, by Hudson, H. (2001). *Materialist metaphysics of the human person* (p. 121). Ithaca, NY: Cornell University Press. Cf. 26–27, 37–38, 101–105, 117, 121–122, 127–128.

Bailey embraces *Property Maximality*. Being a house, for example, is a maximal property, because the proper parts of a house are not themselves houses.[13] *Property Maximality* regarding conscious states,

> *C-Maximality*: Being conscious is maximal *iff* necessarily, for any *x* that is conscious, there is no particular *y*, such that *x* is a proper part of *y*, and *y* is conscious.

Person-pairs cannot be conscious, because they violate *C-Maximality* by attributing conscious persons as proper parts of a person-pair. That each person of the person-pair has its own *what-it-is-like-to-be-a-person* entails that there is no *what-it-is-like-to-be-a-fusion-of-two-persons*.[14]

But why think that consciousness is a maximal property, especially on materialism? Suppose human persons are identical to a physical particular (e.g., a heap, aggregate, system, or a fusion of separable parts). Such particulars have other physical particulars as parts. And most of the intrinsic physical properties of those parts are, or at least could be, intrinsic physical properties of the whole human person, sometimes to a different order of magnitude: having mass, extension (size, shape), charge, powers of gravitational attraction, and so on. None of these properties is maximal, so why think that consciousness is. Given standard naturalism or at least physicalism, there seems to be no nonarbitrary answer to this question.

We think the answer is that subjects of phenomenal consciousness are simple and, thus, could not have separable parts at all, much less conscious ones. The most difficult and unsolved problem for various versions of panpsychism is the combination problem: How can we get a unified consciousness or subject of consciousness by combining tiny conscious particulars? The combination problem is difficult precisely because it is hard to see how a simple subject can be generated from a multiplicity of (attenuatedly) conscious parts. If the combination result in a complex conscious object, then the complexity runs afoul of the simplicity intuition and the unity arguments presented in Chapter 6. It seems that the simplicity/unity issues are the real difficulties at the heart of the combination problem.

Lastly, if consciousness is a maximal property, then what partly determines the fact that any person *S* is conscious is the fact that *S* is not a part of any conscious thing. If that is the case, then any conscious person *S* would become a zombie merely by becoming part of some other conscious thing *S**. Consider Ned Block's Chinese Nation argument. If consciousness is maximal, then when the individual conscious persons of the nation of China realize a function sufficient for consciousness each citizen becomes a zombie, while the nation itself becomes conscious. Likewise, for the clerk in Searle's Chinese room. But that seems absurd.

5.1.2.3 Substance

Eric Yang argues that the following fact explains the Datum. *Substance*: a subject of consciousness must be a substance, and a person-pair isn't a substance.[15] Yang's view entails that conscious subjects must be substances, which explains the *Datum*. We agree that subjects of consciousness must be substances. But not all theories of substance are equal, and not all theories of substance will rival Simplicity as an explanation of the *Datum*. For example, in the Husserlian mereology above, anything without separable parts is a mereologically simple substance. Consequently, any theory that conditions substancehood on having no separable parts as explaining the Datum includes substances with numerous inseparable parts. Examples include versions of hylomorphism, according to which a substance cannot have proper (separable) parts but does have components that are not proper (separable) parts.[16]

---

[13] Bailey follows Ted Sider's treatment of Maximality. Cf. Sider, T. (2001). Maximality and intrinsic properties. *Philosophy and Phenomenological Research*, *63*, 357–364; and Sider, T. (2003). Maximality and microphysical supervenience. *Philosophy and Phenomenological Research*, *66*, 139–149.

[14] Bailey, p. 15.

[15] Yang, E. (2018). Persons, simplicity, and substance. *Philosophical Papers*, *47*(2), 299–311.

So, Yang's *Substance* explanation itself is not necessarily a rival but an instance of Simplicity. Of course, Yang might offer a precise theory of substance that actually does rival Simplicity. The difficulty will be avoiding Simplicity while maintaining a theory of substance that does not either (i) fall into one of Barnett's rival objections or (ii) provide only the Functional Unity of a bundle.

Perhaps some will not be satisfied with our treatment of *Substance*. The objection in the next section is similar to Yang's, although in terms of topological integrity. Plausibly, a substance exhibits a particular kind of topological integrity, so we consider our replies in the next section as also replies to Yang's *Substance* account.

5.1.2.4 Integrity

Rory Madden has advanced the most detailed objection to Barnett.[17] According to Madden, Barnett has misidentified the explanation of the *Datum* as a purely mereological requirement (simplicity), while the actual explanation is topological.[18] Accordingly,

> *Integrity*: Our naïve conception of a conscious subject demands that conscious beings be topologically integrated.

Topological integration is a technical notion, which Madden understands as follows.

> *Topologically Integrated*: *W* is *topologically integrated* $=_{\text{def.}}$ a path can be drawn between any pair of proper parts of *W* without exiting through an exterior surface or boundary of *W*.

Conversely, an object is *Topologically Scattered* just in case one cannot draw such a path between any proper parts of *W*. Prime examples of topologically integrated wholes are, says Madden, the eggs of a hen, an intact cup, and human organisms. Madden's examples of topologically scattered objects include a two-piece bikini, a circle of stones, and a dismantled watch. According to *Integrity*, person-pairs cannot be conscious because person-pairs are more like a two-piece bikini than an intact cup. We have four replies. First, Madden claims that *Integrity* is a better explanation than *Simplicity* because "we have no resistance whatsoever to ascribing consciousness to things which manifestly have proper parts—things like human organisms, cats and dogs—which have feet, heads, eyes, and the rest."[19] However, Madden's claim is radically undermined by the empirical data presented above. This alone means that *Integrity* loses what Madden takes to be its advantage over *Simplicity*.

---

[16] See, e.g., Inman, R. D. (2018). *Substance and the fundamentality of the familiar: A neo-aristotelian mereology*. New York, NY: Routledge. Other examples, plausibly include, Koons, R. (2014). Staunch vs. faint-hearted hylomorphism: Toward an aristotelian account of composition. *Res Philosophica*, *91*(2), 151–177; Marmodoro, A. (2021). Hylomorphic unity. In R. Bliss & J. T. M. Miller (Eds.), *The Routledge handbook of metametaphysics* (pp. 284–299). New York, NY: Routledge; Oderberg, D. S. (2007). *Real essentialism*. New York, NY: Routledge; Rea, M. (2011). Hylomorphism reconditioned. *Philosophical Perspectives*, *25*, 341–358; and Toner, P. (2011). Independence accounts of substance and substantial parts. *Philosophical Studies*, *155*(1), 37–343.

[17] Madden, R. (2015). The naive topology of the conscious subject. *Noûs*, *49*(1), 55–70.

[18] It's worth noting that topological integration is only a variation on the structural explanation of the Datum. As such, one could argue that Madden has failed to offer an alternative explanation of the Datum, unexplored by Barnett. Moreover, one could argue that Madden's version of the structure argument fails against Barnett's treatment of structure. We suspect both responses are promising. Due to lack of space, we leave these arguments unexplored.

Second, Madden is right to point out that integration and scatter are at work in explaining the *Datum*. But it's not obvious that *Integrity* is a rival explanation to *Simplicity*. In our view, topological integration, in the sense relevant to conscious subjects, is a feature of a specific kind of object. These objects are not mereologically simple, but they are integrated. You will recall that mereological simplicity is a formal thesis that requires a material thesis regarding parthood to be metaphysically informative.[20] Neither Barnett nor Madden explain what they mean by "part." For reasons we explained in Chapter 3, we think it safe to assume they mean separable part. This oversight provides a way of arguing that *Integrity* does not rival *Simplicity*.

To be a rival, *Integrity* must be independent from *Simplicity*. Integrity is not independent, or so we argue, because it follows from *mereological simplicity*. To understand this, we need to clarify the nature of inseparable parts. First, an inseparable part cannot be conceived of nor exist as an independent whole separated from the whole of which it is a part. Second, the inseparable parts of a whole permeate each other. In contrast, separable parts are separated in relation to other parts of the whole. Consequently, *W* is simple just in case *W* has no separable parts. If *W* has inseparable parts, *W* is necessarily topologically integrated because a path can be drawn between any pair of *W*'s inseparable parts without exiting through an exterior surface or boundary of *W*. The kind of integrity relevant to explaining why person-pairs are not conscious is grounded in and a feature of *Simplicity*. Consequently, *Integrity* is not a rival explanation of the *Datum*, but a feature of *Simplicity*.

Here is a third reply. *Integrity* seems less and less plausible the more fine-grained one's analysis of the human body or brain becomes. Consider again Madden's examples of topologically integrated wholes: an egg, an intact cup, and the human organism. A cup, for example, has a vast multiplicity of separable parts, especially at the atomic level. Suppose it is true that the intact cup is topologically integrated. It is still the case that the parts of the cup are not topologically integrated as they are *separable parts*. By focusing on the cup's micro-composition, the less plausible it seems that the cup could be conscious, likewise, for the body and brain. The fact that *x* is topologically integrated might explain why we intuitively think of *x* as one object, but it does not explain why we think that *x* is possibly a subject of consciousness. Nor does it explain why we think that *x* is not a subject of consciousness when all we know about *x* is that it is a person-pair.

Consider the following case, modified from a consideration of Madden's:

> *Conscious Cup*: You and I function as the right and left hemispheres of McCartney's brain. While doing so, we fall into a vat at a cup making-factory only to emerge as parts of a cup in such a way that we continue functioning like McCartney's brain.

---

[19] Madden, p. 57.

[20] Recall from Chapter 3. *Separable Part*: *x* is a separable part of some whole *W iff* (i) *x* is a particular, (ii) *x* is a part of *W*, and (iii) it is possible for *x* to exist without being a part of *W*. *Inseparable Part*: *x* is an inseparable part of some whole *W iff* (i) *x* is a particular, (ii) *x* is a part of *W*, and (iii) it is not possible for *x* to exist without being a part of *W*.

Is the cup of which we are parts conscious? If the person-pair we form is not conscious, then likely neither is the cup. What explains the cup's nonconscious state is the fact that the person-pair we form is not conscious. The fact that the cup is topologically integrated is irrelevant.

Lastly, *Integrity*, at least insofar as it is compatible with physicalism, assumes that the body or brain is topologically integrated. But this seems false, at least on standard physicalism (as opposed to staunch hylomorphism or perhaps some version of emergentism). Clouds do not seem topologically integrated. As Peter Unger has pointed out, the boundaries of brains, like the boundaries of clouds, are fuzzy.[21] It isn't clear that you could draw a line between any of the parts of a cloud without exiting through an exterior surface or the boundary of a cloud. And this isn't an epistemological issue but a metaphysical issue brought to light by considering a fine-grained look at clouds. There just isn't a boundary. Similarly, the closer we look at the brain's aggregate structure, the less clear it is that it has boundaries.

Consider, for example, the number of neurons and their synaptic gaps. Neuroscientist Daniel Levitin explains,

> It is difficult to appreciate the complexity of the brain because the numbers are so huge, they go well beyond our everyday experience (unless you are a cosmologist). The average brain consists of one hundred billion (100,000,000,000) neurons.[22]

At chemical synapses, where one neuron communicates with another, neurons do not communicate through bridging channels but by an action potential in the presynaptic cell that leads to the release of a chemical transmitter. Chemical synapses lack structural continuity between the pre- and postsynaptic neurons, resulting in a *synaptic cleft*. This gap between the axon (presynaptic neuron) and the dendrite (postsynaptic neuron) is 20–40 nanometers wide. The enormous number of a brain's separable parts (e.g., neurons) and their separations (e.g., synaptic clefts) presents the extreme aggregation of the brain and its associated lack of holistic unity.

Madden will respond that the problem is not the seeming composite nature of the brain (that it is not mereologically simple) but that the brain appears to be a topologically scattered collection.[23] So, it is Integrity and not Simplicity that explains the *Datum*. Madden makes this argument from the claim that Integrity can make a more discerning diagnosis than Simplicity. But that isn't the case with the Husserlian notion of *Mereological Simplicity*, which is more discerning than *Topological Integrity*. The ontology of separable and inseparable parts specifies the nature of the path between any pair of proper parts of *W*, as well as the nature of *W*'s exterior surface or boundary.

The only truly topologically integrated object is a mereological simple, an object that has no separable parts. *Simplicity* explains the *Datum*. *Simplicity* explains why there is no us in consciousness.

---

[21] Unger, P. (2004). The mental problems of the many. In D. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. 1, pp. 195–222). Oxford: Oxford University Press.

[22] Levitin, D. J. (2006). *This is your brain on music: The science of human obsession* (p. 85). New York, NY: Dutton.

[23] This is Madden's response to Barnett's analogy between the composite natures of the universe and the brain (63).

5.1.3 Summary

Barnett's argument survives the extant objections, and so we should take seriously the thesis that subjects of consciousness are mereologically simple. We think that the story doesn't end here. Barnett's argument pumps an intuition. The intuition itself, however, provides more than an intuition pump, but a strong argument for SD, to which we now turn.

5.2 THE HARD META-PROBLEM OF CONSCIOUSNESS

5.2.1 Intuitions and Rational Seemings

Philosophers differ about the nature of intuitions. On the traditional view, intuitions are cases of (perhaps, defeasible) first-person direct awareness of an intentional object. They are reported by way of the phenomenological use of "seems" or "appears." As Harald Wiltsche has shown, phenomenologists mean something closer to seemings when talking of intuitions.[24] We will therefore talk of intuitions as *rational seemings* or *seemings* for short.[25]

In our view, seemings are ways of being appeared to. We take the concepts expressed in words such as "seems," used in a primitive phenomenological way, to express "intuitions." Stated adverbially, a *seeming* is a way of being appeared to by a purported intentional object. We will have much more to say about this in Chapter 8.

Most philosophers take dualist seemings as significant data in need of an explanation.[26] This is evidenced by recent attention to what David Chalmers calls the *meta-problem of consciousness*: the problem of explaining why we have what he calls *problem intuitions*—intuitions that there is a hard problem of consciousness, including explanatory and metaphysical intuitions, among others.[27] Among these intuitions are dualist seemings. We will explain shortly why we prefer to talk of seemings rather than intuitions.

---

[24] Wiltsche, H. (2015). Intuitions, seemings, and phenomenology. *Teorema*, *34*(3), 57–78. Cf. Bengson, J. (2014). How philosophers use intuition and "intuition." *Philosophical Studies*, *171*(3), 555–576; and Preston, A. (2015). Intuition in analytic philosophy. *Teorema*, *34*(3), 37–56.

[25] See, e.g., Bealer, G. (1998). Intuition and the autonomy of philosophy. In M. DePaul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 207–208). Lanham: Rowman and Littlefield; Bealer, G. (2002). Modal epistemology and the rationalist renaissance. In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility*. New York, NY: Oxford University Press; Bengson, J. (2015). The intellectual given. *Mind*, *124*(495), 707–760; Chudnoff, E. (2011). What intuitions are like. *Philosophy and phenomenological research*, *82*(3), 625–654. Pollock, J. L. (1974). *Knowledge and justification*. Princeton University Press; and Pust, J. (2000). *Intuitions as evidence*. New York, NY: Routledge; and Sosa, E. (2007). Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien*, *74*, 51–67.

[26] Daniel Dennett, often representing the minority view, dismisses the project as being in "the same spirit that economists and psychologists might set out to explain why it is that so many people succumb to the potent illusion that money has intrinsic value." Dennett, D. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness* (p. 178). Cambridge, MA: MIT Press. The two cases are obviously disanalogous. The belief that money has intrinsic value is not a ubiquitous, resilient, pre-theoretical, commonsense belief like SD belief. Elsewhere, Dennett contradicts himself stating that our dispositions to point to experiences of phenomenal consciousness "however unanalyzable, however, intrinsically present that all seems to you, what has to be explained is that it *seems* to, not that it is so." Dennett, D. (2006). You have to give up your intuitions about consciousness. In S. Blackmore (Ed.), *Conversations about consciousness* (p. 88). Oxford: Oxford University Press.

5.2.2 The Easy and Hard Meta-Problems of Consciousness

According to Chalmers, the problem intuitions are dispositions, dispositions to make verbal reports and judgments that, for example, physicalism seems false. But this is mistaken. First, this assumption sneaks in a linguistic reduction. By aiming the problem at dispositions, Chalmers ignores the phenomenological nature of intuitions/seemings in general and dualist seemings themselves. In doing so, Chalmers addresses a much easier task. As we see it, the meta-problem of consciousness can be broken into two problems analogous to Chalmer's distinction between the easy and hard problems of consciousness, as follows.

> *The Easy Meta-Problem of Consciousness*: The difficulty of explaining why we are disposed to report that there is a hard problem of consciousness or that dualism and or SD seem true.
>
> *The Hard Meta-Problem of Consciousness*: The difficulty of explaining why it seems to us that dualism and or SD are true.

Like the *Easy Problem of Consciousness*, that of explaining how things like us can process information in terms of computational or neural mechanisms, the *Easy Meta-Problem of Consciousness* seems solvable mechanistically. On the other hand, most, if not all dispositions have categorical bases/grounds, and Chalmer's linguistic dispositions require such bases/grounds. Why do we have such dispositions? Because we have been appeared to in a phenomenally relevant way.

However, whatever one thinks of the need for a base/ground as part of the solution to the *Easy Problem*, like the *Hard Problem of Consciousness*, that of explaining how things like us are phenomenally conscious, the *Hard Meta-Problem of Consciousness* focuses on explaining a specific kind of state of phenomenal consciousness—*dualist seemings*—that seem to resist mechanistic explanation.

What is more, Chalmers disregards what he calls *self-intuitions*: problem intuitions regarding the subject of consciousness. Instead, only problem intuitions about consciousness itself are taken seriously. Likewise, no attention is given to the cognitive science of commonsense SD. Our point is that dualist seemings have not been dismissed as insignificant by most philosophers whose own views would be more secure by such a dismissal. At the same time, dualist seemings about the subject of consciousness are ignored. As another example, David Papineau ignores both SD seemings as well as SD as an explanation of dualist seemings considering only epiphenomenalism and Russellian monism as competitors with physicalism.[28] While philosophers of mind have taken up the *Hard Problem of Consciousness* they are currently interested only in the *Easy Meta-Problem of Consciousness*.

---

[27] Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, *25*(9–10), 6–61.

[28] Papineau, D. (2020). The problem of consciousness. In U. Kriegel (Ed.), *The Oxford handbook of the philosophy of consciousness* (pp. 14–38). Oxford: Oxford University Press.

As we are about to argue, the extensive evidence that presses the Hard Meta-Problem of Consciousness can no longer be ignored.

5.2.3 The Data: Dualist Seemings

5.2.3.1 Explanatory Gap Seemings

Not only does SD seem true to almost everyone, but physicalism also seems false. In John Perry's words, physicalism "just seems crazy."[29] Likewise, David Papineau explains,

> Indeed I would say that there is a sense in which even professed philosophical physicalists, including myself, cannot fully free themselves from this intuition of distinctness. Of course, we deny dualism in our writings, and take the theoretical arguments against it to be compelling, but when we aren't concentrating, we slip back into thinking of conscious feeling as something extra to the brain.[30]

Intuitions or seemings of distinctness are provoked when considering the prospects of consciousness being or arising from the purely physical or nonconsciousness. Divide all facts about the world into the following:

(P): The complete set of microphysical facts about the universe; and
(Q): Any fact about phenomenal consciousness.

It seems to almost everyone that knowledge of P does not entail or provide knowledge of Q.[31] This produces an *explanatory gap*: the inability to provide or even comprehend a plausible explanation of how Q could depend on P.[32] Thus the hard problem of consciousness. Most think the hard problem is unsolvable, leaving the explanatory gap wide open.[33]

Philosophers draw different inferences from the explanatory gap. Some think it only reflects current scientific limitations.[34] Some think it shows that our cognitive capacities are *in principle* incapable of a solution.[35] Others take the explanatory gap as evidence against physicalism and support property dualism or SD.[36]

---

[29] Perry, J. (2001). *Knowledge, possibility, and consciousness* (p. 4). Cambridge, MA: MIT Press.

[30] Papineau, D. (2008). Explanatory gap and dualist intuitions. In L. Weiskrantz & M. Davies (Eds.), *Frontiers of consciousness* (p. 57). Oxford: Oxford University Press.

[31] Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450; Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, *32*, 127–136; and Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy*, *83*, 291–295.

[32] Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, *64*, 354–361.

[33] See, e.g., Block, N., & Stalnaker, R. (1999). Conceptual analysis, dualism, and the explanatory gap. *The Philosophical Review*, *108*(1), 1. As we saw in Chapter 2, emergent properties, e.g., properties of phenomenal consciousness, are in principle, inexplicable. Accordingly, they are widely taken as brute facts.

[34] Dennett, D. (1991). *Consciousness explained*. New York, NY Back Bay Books; Searle, J. (1992). *The rediscovery of the mind*. Cambridge, MA: MIT Press; and Churchland, P. (1995). *The engine of reason and seat of the soul*. Cambridge, MA: MIT Press.

[35] See, e.g., McGinn, C. (1989). Can we solve the mind-body problem? *Mind*, 98, 349–366.

[36] See, e.g., Foster, J. (1996). *The immaterial self: A defense of the cartesian dualist conception of mind*. London: Routledge, chapter 5; and Robinson, W. (2004). *Understanding phenomenal consciousness*. Cambridge: Cambridge University Press.

Perhaps surprisingly, commonsense dualism and dualist seemings are common among consciousness researchers.[37] Dualist intuitions are even prevalent in the conceptual framework of the neurosciences as a form of explicit substance dualism or a "Cartesian materialism."[38] A strong case has been made that dualism is constantly at work in implicit premises, experiment descriptions, executions, evaluations, and expected conclusions.[39] Explanatory gap seemings are also ubiquitous among nonacademics.[40] The explanatory gap reveals resilient and pervasive dualist seemings.

5.2.3.2 Unity and Simplicity Seemings

It seems to many that subjects of consciousness are metaphysically unified and mereologically simple. As we show in Chapter 6, both dualist seemings go back to Plato and stubbornly return throughout the history of philosophy, cutting across cultures, socioeconomic and historical backgrounds. Moreover, these seemings are recognized as closely associated with SD.

Consequently, we should expect that these dualist intuitions inform contemporary philosophy of mind. David Barnett argues, convincingly in our estimation, that many of the most influential thought experiments in contemporary philosophy of mind are based on the intuition or seeming that we are genuine unities, mereologically simple.[41] Barnett provides the following list:

> *Descartes's Zombie*: Bodies physically identical to ours could lack consciousness.
>
> *Huxley's Explanatory Gap*: There can be no explanation of how states of consciousness arise from interaction among a collection of physical things.
>
> *Putnam's Swarm of Bees*: A swarm of bees could not itself be conscious.
>
> *Block's Miniature Men in the Head*: A collection of tiny men realizing the same functional states as an ordinary brain could not itself be conscious.
>
> *Block's Nation of China*: A collection of ordinary people realizing the same functional states as an ordinary brain could not itself be conscious.

---

[37] For example, an analysis of two surveys at the annual conference of the ASSC (2018 and 2019) revealed that a large majority (84.3%) of active consciousness researchers reject the standard physicalist accounts of Mary in the Black and White Room, instead holding that Mary learns something new when she sees color for the first time. Francken, J., et al. (2021, June 14). An academic survey on theoretical foundations, common assumptions, and the current state of the field of consciousness science. Preprint DOI: https://doi.org/10.31234/osf.io/8mbsk.

[38] This is widely recognized. See, e.g., Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell, see especially 68–81; Rockwell, W. T. (2005). *Neither ghost nor brain*. Cambridge, MA: The MIT Press; 16, 40, 151–52; Uttal, W. R. (2001). *The new phrenology*; and Uttal, W. R. (2004). *Dualism: The original sin of cognitivism*. Mahwah, NJ: Lawrence Erlbaum Associates.

[39] See, e.g., Antonietti, A. (2008). Must psychologists be dualists? In A. Antonietti, A. Corradini, & J. Lowe (Eds.), *Psycho-physical dualism today: An interdisciplinary approach* (pp. 37–68). Lanham, MD: Rowman and Littlefield; and Manzotti, R., & Moderato, P. (2014). Neuroscience: Dualism in disguise. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 81–97). New York: Routledge.

[40] For example, a study by Gottlieb and Lombrozo shows that across a diverse sample, people judge those phenomena tied to privileged access and subjective experience are difficult if not impossible to explain scientifically. See Gottlieb, S., & Lombrozo, T. (2018). Can science explain the human mind? Intuitive judgments about the limits of science. *Psychological Science*, 29(1), 121–130.

[41] See Barnett, D. (2008). The simplicity intuition and its hidden influence on philosophy of mind. *Noûs*, *42*(2), 308–335.

> *Searle's Chinese Room*: A system comprising a person who does not understand Chinese and a written set of rules could not itself understand Chinese.
>
> *Unger's Zuboffian Brain Separation*: A collection of widely scattered neurons could not itself be conscious.

Some of these thought experiments are present throughout the history of philosophy.[42] That these thought experiments have been recently recreated without reference to their origin is most charitably explained by the fact that dualist intuitions are epistemically strong and ubiquitous. This fact also provides at least a *prima facie* defeater for modal skeptics about these and other thought experiments that support dualist seemings. That our appeal to dualist intuitions for why these thought experiments seem so obviously correct must be explained.

5.2.3.3 The Cognitive Science of Commonsense Substance Dualism

Numerous cognitive science studies demonstrate the widespread commonsense belief in SD. After surveying numerous developmental psychology studies, Henry Wellman concludes that "young children are dualists: knowledgeable of mental states and entities as ontologically different from physical objects and real [nonimaginary] events."[43] Neuropsychologist Paul Broks calls belief in dualism a "primordial intuition." The evidence confirms that we are "natural born soul makers, adept at extracting unobservable minds from the behavior of observable bodies, including their own."[44] Cognitive psychologist, Paul Bloom, arrives at the same conclusion in early social cognition research.[45] Referencing multiple studies, Bloom argues that infants instinctually imitate the facial expressions of adults,[46] have expectations about social interactions with people,[47] and display different expectations for objects than people, in that they do not appear surprised when people violate object principles but are surprised when objects violate those principles.[48] This data shows, argues Bloom, that infants can distinguish between bodies and persons and only attribute intentions (mental states) to persons.

---

[42] In Chapter 6 we develop this point in detail.

[43] Wellman, H. (1990). *The child's theory of mind* (p. 50). Cambridge, MA: MIT Press. We owe this reference to Stewart Goetz and Mark Baker.

[44] Broks, P. (2005). Out of mind. *Prospect*, *109*(April), 1; as quoted in Humphrey, N. *Soul dust: The magic of consciousness* (p. 95). Princeton: Princeton University Press.

[45] Cf. Bloom, P. (2004). *Descartes's baby: How the science of child development explains what makes us human*. New York: Basic Books; idem. (2007). Religion is natural. *Developmental Science*, *10*, 147–151; and Bloom. (2009). Religious belief as an evolutionary accident. In Murray, M., & Schloss, J. (Ed.), *The believing primate: Scientific, philosophical, and theological reflections on the origin of religion* (pp. 118–27). New York: Oxford University Press.

[46] Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 83–99.

[47] Rochat, P., & Striano, T. (1999). Social-cognitive development in the first year. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first year of life* (pp. 3–34). Mahwah, NJ: Lawrence Earlbaum.

[48] Kuhlmeier, V. A., Bloom, P., & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition*, *94*(1), 95–103.

Wellman has conducted several landmark studies acknowledged as producing significant and well-established findings.[49] The first experiment involved 72 preschoolers, while the second involved 60 preschoolers. These children were asked, for example, if they could perform various tasks without a brain and separately without a mind. These tasks included mental acts (thinking, remembering), perception (seeing, hearing), feelings (interests, happiness), voluntary action (walking, talking), and involuntary behavior (breathing, sneezing). The results of these studies show the following:

(a) Children appeal to the commonsense criteria of behavioral sensory evidence (visibility, tangibility) to distinguish mental entities from prototypic physical objects (e.g., a thought about a dog from an actual dog).
(b) Children distinguish physical entities from intangible or invisible mental entities (e.g., smoke from sounds).
(c) Children understand that mental entities are peculiarly private.
(d) Children judged mental representational entities (e.g., mental images) as quite different from physical representation (e.g., photographs).
(e) Children responded correctly to questions about mental entities with unspecified real referents (a thought about a dog), as well as about mental entities with specified referents (a mental image of a cup).
(f) Children recognize that mental entities can be about nonexistent states of affairs.

These children exhibit "a solid and articulate understanding of the fundamental distinction between mental entities and physical objects."[50] Wellman concludes, "Our data argues clearly that the first understanding of mind is one of ontological dualism ..."[51] Similar studies show that young children have a commonsense conception of agents as compounds of body and soul.[52]

More recently, developmental psychologists Jesse Bering and David Bjorklund conducted experiments where a puppet show depicts an alligator eating a mouse.[53] This was shown to three different age groups: kindergartners (ages 3–6), elementary-schoolers (ages 10–12), and adults (ages 18–20). After the puppet show, participants were asked a set of questions about the mouse who died. These questions examined six aspects of death: biological, psychobiological, perceptual, desire, emotional, and epistemic. The study revealed a discontinuity among the age groups regarding every question set except the epistemic question set. The following year, a similar experiment compared religiously and secularly schooled children.[54] The results were similar. According to Bering and Bjorklund, the data suggest that belief in an afterlife is not acquired through social learning. Justin Barret explains that SD beliefs "are among the most widespread supernatural belief. Even in places that claim to be atheistic."[55]

---

[49] See, e.g., Wellman, H. M., & Estes, D. (1986). Early understanding of mental entities: A re-examination of childhood realism. *Child Development*, *57*, 910–923; Wellman, H. M., & Hickling, A. K. (1994). The minds "I": Children's conception of the mind as an active agent. *Child Development*, *65*, 1564–1580; Estes, D., Wellman, H. M., & Woolley, J. (1989). Children's understanding of mental phenomena. In H. Reese (Ed.), *Advances in child development and behavior* (p. 22). New York: Academic Press; and Wellman, *The child's theory of mind*.

[50] Wellman, *The child's theory of mind*, p. 60.

[51] Ibid., p. 59.

[52] See, e.g., Spelke, E. S., et al. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44–78). Oxford: Clarendon Press; and Saxe, R., Tzelnic, R., & Carey, S. (2006). Five-month-old infants know humans are solid, like inanimate objects. *Cognition*, *101*(1), B1–B8.

[53] Bering, J., & Bjorklund, D. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, *40*(2), 217–233.

Dualist seemings are also common among neuroscientists. An analysis of two surveys at the annual conference of the ASSC (2018 and 2019) revealed that a large majority (84.3%) of active consciousness researchers reject the standard physicalist accounts of Mary in the Black and White Room, instead holding that Mary learns something new when she sees color for the first time.[56] Dualist seemings are even prevalent in the conceptual framework of the neurosciences. as an explicit substance dualism or a "Cartesian materialism."[57] A strong case has been made that dualism is constantly at work in implicit premises, experiment descriptions, executions, evaluations, and expected conclusions.[58]

A massive amount of data, more than we can mention, strongly suggests that SD is a naturally formed belief, that dualist seemings are ubiquitous and strong.[59]

5.2.3.4 Eighteen Distinct Dualist Seemings

There are many apparently unrelated items of commonsense self-knowledge, usually expressed in the first-person. It strongly seems to me that,

(i) *Distinctness*: States of consciousness are not physical, although brain states are physical.
(ii) *Intentionality*: My thoughts are, at least most of the time, about real things.
(iii) *Subject Necessity*: A state of consciousness cannot exist without a subject of that state, but a brain state can exist without a subject of consciousness existing.
(iv) *Mine-ness*: My states of consciousness are uniquely and *de re* necessarily mine, but my brain states could be states of someone else. I have no difficulty determining which mental states are mine. Indeed, I am directly aware (1) of my mental states,

---

[54] Bering, J., Blasi, C. H., & Bjorklund, D. (2005). The development of "afterlife" beliefs in religiously and secularly schooled children. *British Journal of Developmental Psychology*, *23*, 587–607.

[55] Barret, J. L. (2012). *Born believers: The science of children's religious belief* (p. 215). New York: Free Press.

[56] Francken, et al. (2022).

[57] This is widely recognized. Cf. Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell, see especially 68–81; Rockwell, W. T. (2005). *Neither ghost nor brain*. Cambridge, MA: The MIT Press; 16, 40, 151–152; Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: The MIT Press; and Uttal, W. R. (2004). *Dualism: The original sin of cognitivism*. Mahwah, NJ: Lawrence Erlbaum Associates.

[58] See, e.g., Antonietti, A. (2008). Must psychologists be dualists? In A. Antonietti, A. Corradini, & J. Lowe (Eds.), *Psycho-physical dualism today: An interdisciplinary approach* (pp. 37–68). Lanham, MD: Rowman and Littlefield; and Manzotti & Moderato, Neuroscience: Dualism in disguise, 81–97.

[59] See, e.g., Forstmann, M., & Burgmer, P. (2017). Antecedents, manifestations, and consequences of belief in mind-body dualism. In C. Zedelius, B. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health*. New York: Springer; Forstmann, M., & Burgmer, P. (2015). Adults are intuitive mind-body dualists. *Journal of Experimental Psychology*, *144*, 222–235; and Forstmann, M., Burgmer, P., & Mussweiler, T. (2012). The mind is willing, but the flesh is weak: The effects of mind-body dualism on health behavior. *Psychological Science*, *23*, 1239–1245.

(2) of their belonging to me, and (3) of their being "inside" of me. According to Colin McGinn, materialism seems implausible because we naturally form "the idea that the mind is a *substance*, that mental states are states of a mental substance."[60]

(v) *Mental Privacy*: I have unique epistemic privacy with my mind.

(vi) *Embodied Privacy*: I am aware of having epistemic, proprioceptive immediacy with respect to my entire body.

(vii) *Neural Contingency*: The brain state I am in right now could obtain without the conscious state I am in right now obtaining.

(viii) *Zombie*: An exact physical duplicate of me could behave just as I do but have no conscious experiences.

(ix) *Mental Powers Essentialism*: Some of my mental powers/capacities are essential to my identity, but my physical powers are not.

(x) *Phenomenal Unity*: I have a single phenomenally unified state of consciousness. My states of consciousness (my thoughts, sensations, and so forth) are united into a single consciousness state, but my brain states are not.

(xi) *Subject Unity*: I am a single unified subject of consciousness, not a collection of conscious subjects.

(xii) *Simple Self*: Pairs of subjects of consciousness cannot themselves be a subject of consciousness. Siamese twins involve two, not three subjects of experience. When two people shake hands, a new subject of experience does not come-to-be. We are inclined to think that this is because subjects of experience cannot be composites. Thus, in whatever relational structure, pairs or larger groups of anything cannot be a conscious subject of experience.

(xiii) *Endurance*: I am the same self that lives through and owns each successive phenomenal experience over time. As I walk toward my kitchen, this seems. In the middle of the sequence, I have strong intuitions of having had earlier experiences, of currently having a particular experience, and of being about to have an anticipated experience.

(xiv) *Physical Persistence*: I can continue to exist even though my physical parts are replaced, could be completely replaced, or could have different in a possible world in which I myself exist.

> Regarding issues in personal identity, Frank Jackson acknowledges: "I take it that our folk conception of personal identity is Cartesian in character—in particular, we regard the question of whether I will be tortured tomorrow as separable from the question of whether someone with any amount of continuity—psychological, bodily, neurophysiological, and so on and so forth—with me today will be tortured."[61]

(xv) *Psychological Persistence*: I can continue to exist even if my memories, character, and personality traits instantly change, and that they all remain the same does not ensure that I continue to exist.

> Others could satisfy psychological criteria of personal identity and not be me, and I could fail to satisfy them and still be me. Various causal chain analyses regarding problems with psychological criteria are (a) still neither necessary nor sufficient for personal identity and (b) part of a dialectic involving stepwise causal-chain analyses, counterexamples, new analyses,

---

[60] McGinn, C. (1989). *Mental content* (p. 15). Oxford: Basil Blackwell.

[61] Jackson, F. (1998). *From metaphysics to ethics* (p. 45). Oxford: Clarendon.

> new counterexamples, and so forth that are guided by a concept of the self as a simple, spiritual substance. And I have strong intuitions that appeal to immanent causation presuppose sameness of the self over time and, thus, cannot constitute it.

(xvi) *Indivisibility*: I cannot exist in degrees but am indivisible.

> If I lose, say, half of my body or brain, I am not thereby a half of a person. I am an all-or-nothing kind of thing—either I am present, or I am not. In some brain operations and cases of Dandy-Walker Syndrome, well over half of the brain is absent (in many cases of Dandy-Walker Syndrome, only 10% of the brain is present.) But while the person may lose some functioning, we have strong intuitions that a whole person is present.

(xvii) *Disembodiment*: I can exist without a body.

(xviii) *NDE*: (a) Near-death cases are (metaphysically) possible and (b) whether or not they are real is a function of the eyewitness (and related) evidence and not a function of the laws of physics/chemistry or the survival of my brain and body.

Plausibly, commonsense beliefs among these seemings correlate (i)–(x) with at least mere property dualism, while (xi)–(xviii) correlate with substance dualism. Moreover, adopting property or substance dualism exhibits the significant explanatory virtue of taking what appears to be a list of commonsense beliefs and unifying/simplifying them around the relevant form of dualism.

5.2.3.5 Artificial Intelligence Test Seemings

A further example of the ubiquity of dualist seemings is their role in artificial intelligence research. Susan Schneider and Edwin Turner have developed the *AI Consciousness Test* (ACT), which asks an AI about the possibility of its survival through body swapping, out-of-body experiences, and the afterlife.[62] An AI that gives even one satisfactory answer to any ACT is sufficient for passing the test. Why think these questions are relevant indicators of consciousness? Schneider explains:

> Notice that normal adults can quickly and readily grasp certain concepts based on the quality of felt consciousness—that is, the way it feels, from the inside, to experience the world … the point is that we can imagine them, at least in broad strokes, because we are conscious beings.[63]

Three things are important for our purposes. First, ACT assumes, if it is to have explanatory power, that at least some dualist seemings are ubiquitous among human persons. Second, dualist seemings are not fringe but a significant indicator of consciousness. Third, dualist seemings are the result of experiencing one's own consciousness.

5.2.3.6 Three Explanatory Conditions of Dualist Seemings

Three features of dualist seemings from the above data must be explained:

> *Diversity*: Dualist seemings are numerous and do not stand in any obvious explanatory connection to each other.

---

[62] For a short list of ACT sample questions, see Schneider, S. (2019). *Artificially you: AI and the future of your mind* (p. 55). Princeton, NJ: Princeton University Press.

[63] Schneider, S. (2020). How to catch an AI zombie: Testing for consciousness in machines. In S. Matthew Liao (Ed.), *Ethics of artificial intelligence* (p. 443). Oxford: Oxford University Press.

> *Ubiquity*: Dualist seemings are pervasive across every known period of human history, human cultures, and stages of psychological development.
>
> *Strength*: Dualist seemings are psychologically and phenomenologically persuasive at a base commonsense level.

Among other things, *Diversity* requires that any plausible explanation cannot reduce dualist seemings to one seeming under different descriptions. Diversity also means that the focus on explanatory gap intuitions fails to recognize the wide range of data. Explaining one dualist seeming does not entail explaining all of them. Consequently, as we just noted, a successful explanation connects or unifies the total set of dualist seemings in an explanatorily significant way.

Consciousness studies rarely discuss the empirical details of *Ubiquity*. Dualist seemings, especially SD seemings, extend much farther back in human history than most recognize, likely to the beginning of human civilization.[64] Moreover, SD seemings cut across cultural and religious divides. Not only prevalent in ancient Greece and ancient Egypt,[65] eastern religions, including Hinduism and some forms of Buddhism, express commonsense SD, as do native North Americans.[66] A large-scale quantitative examination of mind and body concepts in pre-Qin Chinese literature (pre-221 BCE) revealed both the universality of pre-theoretical SD belief as well the intellectual development of SD in early China.[67] Contrary to the standard mythology, SD seemings are even prevalent in self-professing atheist cultures.[68] As already mentioned, the prevailing view among cognitive scientists is that SD seemings are prevalent not just among adolescents and adults but are present in infancy.

Strength: dualist seemings are phenomenologically and epistemically persuasive at a base commonsense level. According to Nagel, no amount of empirical evidence will make it possible for most to abandon dualism for physicalism.[69] Dualist seemings do not yield to a revisionary physicalist ontology. Yet, SD objectors suggest we dismiss these seemings as insignificant or misleading. However, it isn't always rational to abandon the obvious. Thomas Kelly explains:

> But to the extent that a change in view inspired by philosophy is a reasonable one, it can be represented as the product of reasoning, reasoning which proceeds from premises ... In order for such a philosophy-inspired change in view to be reasonable, it is of course not enough that the premises from which the reasoning proceeds are in fact true; what is required is that they are reasonably believed to be true.[70]

---

[64] See, e.g., Martin, R., & Barresi, J. (2006). *The rise and fall of soul and self: An intellectual history of personal identity* (p. 290). New York, NY: Columbia University Press.

[65] See, e.g., Herodotus (ca. 425 BCE), *Historiae* 2, 123, 2.

[66] See, e.g., Bremmer, J. (1983). *The early Greek concept of the soul.* Princeton, NJ: Princeton University Press, chapter 1; Roazzi, M., Nyhof, M., & Johnson, C. (2013). Mind, soul and spirit: Conceptions of immaterial identity in different cultures. *International Journal for the Psychology of Religion*, *23*, 75–86.

[67] Slingerland, E., & Chudek, M. (2011). The prevalence of mind-body dualism in early China. *Cognitive Science*, *35*(5), 997–1007.

[68] See, e.g., Barret, *Born believers*, p. 215; and Bering, J., Blasi, C. H., & Bjorklund, D. (2005). The development of "afterlife" beliefs in religiously and secularly schooled children. *British Journal of Developmental Psychology*, *23*, 587–607.

[69] Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450.

[70] Kelly, T. (2008). Common sense as evidence: Against revisionary ontology and skepticism. *Midwest Studies in Philosophy, 32*, 59.

Such beliefs are called Moorean facts. In the words of David Lewis, a Moorean fact is “one of those things that we know better than we know the premises of any philosophical argument to the contrary.”[71] Dualist seemings are so strong that belief in SD qualifies as a Moorean fact.

Walter Hopp observes, "For many … approaches to the mind, conscious experiences are theoretical postulates postulated to explain certain observable data. For the phenomenologist, they are the data.”[72] Hopp is correct, and responses to dualist seemings by non-dualists reveal as much. Any theory of mind must plausibly explain the dualist seemings data, which includes explaining *Diversity*, *Ubiquity*, and *Strength*. If one is going to reject dualist seemings, then one needs to explain these features plausibly. “Even if materialism isn’t mistaken,” says Papineau, “its defenders owe an explanation of why it should *seem* so mistaken.”[73] This is no small or insignificant problem. As Jerry Fodor has remarked, if the ordinary person’s understanding of the mind should turn out to be seriously mistaken, it would be “the greatest intellectual catastrophe in the history of our species.”[74] In what follows, we attempt to avoid this catastrophe.

5.3 A DIRECT SELF-AWARENESS ARGUMENT FOR SUBSTANCE DUALISM

5.3.1 DSA: The Direct Self-Awareness Account of Dualist Seemings

In what follows, we will argue that the following SD thesis best explains dualist seemings in purported acts of self-awareness.

> *Direct Self-Awareness Account* (DSA): Facts about dualist seemings are grounded in facts about being directly aware of oneself, which reveals oneself as a mereologically simple, immaterial substance.

Facts about dualist seemings, including *Diversity*, *Ubiquity*, and *Strength*, are grounded in facts about one being directly aware of oneself.

The veracity of self-awareness or self-consciousness is a longstanding view. Aristotle, for example, claims that a person must, while perceiving anything, also perceive his/her own existence (*De Sensu* 7.448a). Yet, self-consciousness presents serious problems for naturalist theories of consciousness,[75] acquaintance,[76] and intentionality.[77] There are many other related problems largely unacknowledged by SD dissenters.[78] Others argue that various metaphysical features of self-knowledge or self-consciousness provide compelling evidence for substance dualism.[79]

---

[71] Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, *74*(4), 549.

[72] Hopp, W. (2011). *Perception and knowledge* (p. 4). Cambridge: Cambridge University Press.

[73] Papineau, D. (2002). *Thinking about consciousness* (p. 162). Oxford: Oxford University Press.

[74] Fodor, J. (1987). *Psychosemantics* (p. xii). Cambridge, MA: Bradford Books/MIT Press.

[75] See, e.g., Bealer, G. (1997). Self-consciousness. *The Philosophical Review*, *106*(1), 69–117.

[76] Brie Gertler observes, “the process of grasping an introspected property through acquaintance may be difficult to accommodate within a broadly naturalistic view of the mind, since many prominent naturalist accounts of mentality construe mental processes in purely causal terms.” Gertler, B. (2011). *Self-knowledge* (p. 126). New York, NY: Routledge. Few have even attempted a physicalist account of acquaintance. See, e.g., Balog, K. (2012). Acquaintance and the mind–body problem. In S. Gozzano & C. Hill (Eds.), *New perspectives on type identity: The mental and the physical* (pp. 16–42). Cambridge: Cambridge University Press.

That said, most philosophers, including naturalists, do not reject self-awareness or self-knowledge. In the words of Daniel Dennett,

> I think the reason that we find consciousness so hard is that we have evolved a certain capacity for self-knowledge, a certain access to ourselves which gives us subjective experience—which gives us a way of looking out at the world from where we are. And this just turns out to be very hard to understand.[80]

Our argument centers on explaining why we have dualist seemings and providing a unification of them. In this account, I am acquainted with myself as a substance with features revealed by directing my intentionality on myself. Dualist seemings are explained by the fact that we are directly acquainted with ourselves, such that SD seems true, and physicalism seems false. To explain this account, we must clarify what we take to be the ontological structure of knowledge by acquaintance, intentionality, and self-presenting properties.[81]

---

[77] For example, certain functionalist reductions of intentionality entail *mental holism*: that a specific mental state is individuated by its relation to one's entire psychology. How anyone could know what concept one was entertaining at a given moment or grasp its intentional features becomes quite opaque for such a view. We think similar difficulties confront attempts to identify beliefs (or thoughts) with indicators (a.k.a. "indicator meanings" or "representations") that satisfy some functional or causal criterion. We suggest that the nature of fulfillment structures provides a test that any view of intentionality must meet on pain of being disregarded as inadequate. While we cannot argue the point here, we believe that (at least) property dualism for the relevant mental states is the only adequate solution that meets this criterion. For more detailed problems see, e.g., Haldane, J. J. (1989). Naturalism and the problem of intentionality. *Inquiry*, *32*(3), 305–322; Willard, D., & Rickabaugh, B. (2018). Intentionality contra physicalism: On the mind's independence from the body. *Philosophia Christi*, *20*(2), 497–515; Meixner, U. (2006). Classical intentionality. *Erkenntnis*, *65*, 25–45. Cf. Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417–424.

[78] Another noteworthy yet overlooked problem is a plausible, *prima facie* conflict between physicalism in philosophy of mind and internalism in epistemology, as the internalist and externalist have desperate theories about mental processes, especially those relating to epistemic justification. As a further example, Michael Rea argues that naturalism and arguments against dualism undermine belief in the existence of other minds. See Rea, M. C. (2002). *World without design: The ontological consequences of naturalism* (pp. 166–168). Oxford: Oxford University Press. We will have more to say about these and related issues in Chapter 11 where we discuss future research programs for substance dualism.

[79] See, e.g., Lund, D. (2005). *The conscious self: The immaterial center of subjective states*. Amherst, NY: Humanity Books; and Goetz, S. (2001). Modal dualism: A critique. In K. Corcoran (Ed.), *Body, soul, and survival* (pp. 89–103). Ithaca, NY: Cornell University Press.

[80] Dennett, You have to give up your intuitions about consciousness, pp. 79–80.

[81] Others have defended SD from adverbial views of intentionality or phenomenal states. Although, these arguments are quite different from ours. See, e.g., Zimmerman, D. (2011). From experience to experiencer. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis; investigations into the existence of the soul* (pp. 168–196). London: Continuum.

5.3.2 Knowledge by Acquaintance

If I am acquainted with something, then I am directly aware of it; *epistemically direct*—my awareness does not rely on an awareness of anything else, and *metaphysically direct*—my awareness is not mediated by or through anything. Not even concepts stand between subjects and the world as it is, and they certainly don't create the world.[82] Even in cases of seeing an entity as *so-and-so*, it does not follow that my seeing the entity must be as *so-and-so*. My seeing an entity as *so-and-so* can be revised by virtue of closer examination.[83]

In the past decade or two, knowledge by acquaintance has made a comeback.[84] In modern philosophy, the prominence of knowledge by acquaintance is usually associated with Bertrand Russell.[85] However, according to some, it largely passed off the scene due to Wittgenstein's private language argument[86] and Sellars' attack on the "Myth of the Given."[87] However, we think a third factor was equally prominent, viz., the concern to reduce phenomenal consciousness—especially sensory states—to linguistic behavior. This factor was a key concern among early advocates—e.g., J. J. C. Smart and D. M. Armstrong—of type identity theory. The reduction was accomplished in two steps (illustrated by visual sensations):[88]

> (Step 1) *Semantic Reduction*: Subject S has a visual sensation of P $=_{\text{def.}}$ Subject S has an inclination to believe that P that is caused in the right way by S's visual sensory organs.
> (Step 2) *Linguistic Reduction*: Subject S has an inclination to believe that P that is caused in the right way by S's visual sensory organs $\geq_L$ under the right circumstances C, S is disposed to engage in behavior, especially verbal reports, that express the ability to use "P" with the proper know-how according to S's linguistic community.

Russell himself took states of acquaintance to be irreducible and conceptually or epistemically fundamental relative to knowledge by description. He also held that there were two critical aspects of knowledge by acquaintance: an epistemic and a metaphysical one. This has raised the following methodological problem: Should knowledge by acquaintance be approached primarily as an epistemic issue or as an ontological one in philosophy of mind? According to the latter, it seems that the essential nature of an act of knowledge by acquaintance can only be described phenomenologically as a special, irreducible state of consciousness. As many (including us) understand a case of knowledge by acquaintance, it involves an experientially pure, *de re*, direct presentation/awareness of an intentional object to the subject and has a unique kind of intrinsic *what-it-is-like-for-S* quality. This presents a problem for physicalists.[89]

---

[82] See Willard, D. (1999). How concepts relate the mind to its objects: The "God's eye view" vindicated? *Philosophia Christ*, *2*(1)(2), 5–20.

[83] Wolterstorff, N. (2011). The role of concepts in our access to reality. In B. L. Gordon & W. A. Dembski (Eds.), *The nature of nature: Examining the role of naturalism in science* (pp. 266–267). Wilmington, DE: ISI Books.

[84] See Raleigh, T. (2019). Introduction: The recent renaissance of acquaintance. In J. Knowles & T. Raleigh (Eds.), *Acquaintance: New Essays* (pp. 1–33). Oxford: Oxford University Press.

[85] Russell, B. (1911). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society 11*, 108–128; *Problems of philosophy* (2nd ed.). Oxford: Oxford University Press, chapter 5; On the nature of acquaintance. (1914). *Monist, 24*, 435–453.

[86] Wittgenstein, L. (1953). *Philosophical investigations* (G.E.M. Anscombe & R. Rhees, Eds., G.E.M. Anscombe, Trans.). Oxford: Blackwell.

[87] Sellars, W. (1956). Empiricism and the philosophy of mind. In H. Feigl & M. Scriven (Eds.), *Minnesota studies in the philosophy of science* (Vol. I, pp. 127–196). Minneapolis, MN: University of Minnesota Press.

[88] For a discussion and critique of Armstrong's employment of this reduction, see Swinburne, R. (1997). *The evolution of the soul* (rev. ed., pp. 34–39, 42–44). Oxford: Clarendon Press.

We differ in various ways from Russell and many after him. We take intentionality to be a monadic property of phenomenal states, not a relation. We take the properties that constitute a state of knowledge by acquaintance to be self-presenting and, specifically, we do not limit the intentional objects of such states to conscious entities within the subject, nor do we limit the intentional objects of such states to sense-perceptual ones.

The link between knowledge by acquaintance and self-awareness is straightforward. Self-awareness or self-knowledge is a kind of acquaintance.[90] Even Dennett admits that self-knowledge and subjective experience are facts of our experience that are "not closely paralleled with anything we know about anything else."[91] In fact, in the received view, being an agent or a subject of consciousness requires that one is epistemically acquainted with at least some features of oneself in a way that no other person can.[92] Despite the acknowledged tension with naturalism, acquaintance has become difficult for many to deny.[93]

---

[89] See, e.g., Adam Pautz, "The Significance Argument for the Irreducibility of Consciousness," *Philosophical Perspectives* 31 (2017): 349–407.

[90] See, e.g., Balog, K. (2012). Acquaintance and the mind-body problem. In S. Gozzano & C. Hill (Eds.), *New perspectives on type identity: The mental and the physical* (pp. 16–43). New York, NY: Cambridge University Press; Balog, K. (2019). Hard, harder, hardest. In A. Sullivan (Eds.), *Sensations, thoughts, and language: Essays in honor of Brian Loar* (pp. 265–289). New York, NY: Routledge; Coleman, S. (2019). Natural acquaintance. In J. Knowles & T. Raleigh (Eds.), *Acquaintance: New essays* (pp. 49–74). Oxford, UK: Oxford University Press; Duncan, M. (2015). We are acquainted with ourselves. *Philosophical Studies*, *172*(9), 2531–2549; Duncan, M. (2017). Two Russellian arguments for acquaintance. *Australasian Journal of Philosophy, 95*(3), 461–474; Duncan, M. (2018). Subjectivity as self-acquaintance. *Journal of Consciousness Studies, 25*(3–4), 88–111; Duncan, M. (2020). Knowledge of things. *Synthese*, *197*(8), 3559–3592; Fumerton, R. (2019). Acquaintance: The foundation of knowledge and thought. In Knowles & Raleigh (pp. 245–259); Gertler, B. (2012). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 89–123). New York, NY: Oxford University Press; Goff, P. (2015). Real acquaintance and physicalism. In P. Coates & S. Coleman (Eds.), *Phenomenal qualities: Sense, perception, and consciousness* (pp. 121–146). Oxford, UK: Oxford University Press; Levine, J. (2019). Acquaintance is consciousness and consciousness is acquaintance. In Knowles and Raleigh (pp. 33–48); Nida-Rümelin, M. (2007). Grasping phenomenal properties. In T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge* (pp. 307–336). New York, NY: Oxford University Press; Pitt, D. (2019). Acquaintance and phenomenal concepts. In S. Coleman (Ed.), *The knowledge argument* (pp. 87–101). Cambridge, UK: Cambridge University Press; and Strawson, G. (2017). *The subject of experience*. Oxford: Oxford University Press.

[91] Dennett, D. You have to give up your intuitions about consciousness, p. 80.

[92] See, e.g., Gertler, *Self-knowledge*, pp. 2–4.

[93] Matt Duncan lists 55 contemporary philosophers who defend a substantive theory of acquaintance. See Duncan, M. (2021). Acquaintance. *Philosophy Compass*, *16*(e12727). (Online first). https://doi.org/10.1111/phc3.12727, fn. 9.

5.3.3 Monadic Intentionality

In our view, intentionality is a *monadic property*—a property instantiated by one individual—and thus, not a polyadic or relational property that is exemplified by most, if not all, mental states, which consists of a primitive *consciousness-of*, *ofness*, *aboutness*, *directedness* toward an object or possible object for a subject. We, therefore, accept two largely uncontested features of intentionality:

> *Primitive*: P is an intentional state *iff* P consists of a primitive *consciousness-of*, *ofness*, *aboutness*, *directedness* toward an object or possible object O.
>
> *Mentality*: P is an intentional state *iff* P is exemplified by a mental state M.[94]

We will assume both are true.

The monadic property view of intentionality often adopts an adverbial theory of awareness.[95] Plausibly Franz Brentano,[96] but certainly Edmund Husserl, held a non-relational—monadic property—theory of intentionality,[97] which is rightly understood as an adverbial view.[98] Curt Ducasse and Roderick Chisholm are well-known defenders of the adverbial monadic property view of intentionality, which we will follow.[99]

In adverbialism, the property <being intentionally directed toward F> becomes <being intentionally directed F-wise> and is expressed as "being appeared to F-wise." Consequently, statements such as, "I am being appeared to F-wise" tell us how or the way the subject is sensing.[100] Moreover, that I am thinking of an object does not itself entail a relation to that object. This is certainly true when thinking of things that don't exist (e.g., Marvel Comics' Iron Man). Rather, I would be *thinking about nothing*

---

[94] It is important to note that this condition is consistent with the existence of some mental states (e.g., a pain) that fail to have intentionality. However, if there are such states, they seem to entail (for those subjects who can have second-order mental states) associated intentional states (e.g., a second-order state of being aware of the first-order pain).

[95] Not all defenders of the monadic property or non-relational intentionality view adopt adverbialism. For example, Tim Crane distinguishes between the intentionality of a thought (non-relational) and the reference of a thought. Although all thoughts are intentional, some thoughts fail to refer (e.g., when one thinks about a nonexistent object). See Crane, T. (2001). Intentional objects. *Ratio*, *14*, 336–349; and Crane, T. (2013). *The objects of thought*. New York: Oxford University Press.

[96] On a plausible reading, Brentano, in his later works, adopted a non-relational view of intentionality. See, e.g., Brentano, F. (1995). *Psychology from an empirical standpoint* (O. Kraus, Ed., L. L. McAlister, Trans.) (pp. 271–272). London: Routledge. First published in German, 1874.

[97] See, e.g., Husserl, E. (1900–01/1970). *Logical investigation* (J. N. Findlay, Trans.) (Vol. 2). London: Routledge and Kegan Paul, V, §11 and § 12, 557–560. David Woodruff Smith and Ronald McIntyre also refer to Husserl's 1894 essay on "Intentional Objects" and his 1896 review of Twardowski.

[98] See Smith, D. W., & McIntyre, R. (1982). *Husserl and intentionality: A study of mind, meaning, and language* (pp. 141–142). Dordrecht: D. Reidel Publishing.

[99] See, e.g., Chisholm, R. (1957). *Perceiving: A philosophical study* (pp. 168–185). Ithaca, NY: Cornell University Press; and Ducasse, C. J. (1968). Moore's refutation of idealism. In P. A. Schilpp (Ed.), *The philosophy of G. E. Moore* (pp. 232–233). La Salle, IL: Open Court.

[100] Chisholm, R. (1969). On the observability of the self. *Philosophy and Phenomenological Research*, *30*(1), 16.

Iron-Man-ly.[101] Intentional acts are, thus, not only given to the subject but are modifications of the subject.[102]

Taking the above theses together entails the following model of intentionality, which we find quite attractive:

> *Monadic Intentionality*: P is a state of intentionality *iff* (i) P is a *monadic property* (ii) P is exemplified by a mental state M, (iii) M is a mode of a subject S, (iv) M consists of a primitive *consciousness-of*, *ofness*, *aboutness*, *directedness* toward an object or possible object, and (v) M is given to S.

That intentionality is both necessarily given to S, and a modification of S is not trivial. My consciousness of <this plant in front of me> has the phenomenological feature of being given to me. As Meixner and others insightfully point out, the subject of intentionality is often absent in definitions of intentionality.[103] This evidences a serious failure of contemporary philosophy of mind to either eliminate the self-subject or make the self-subject epiphenomenal.

5.3.4 Givenness and Directedness

First, intentionality has both a directed-at-ness and given-to-ness; a directed-at-ness of an intentional object and a given-to-S-ness. It is uncontroversial that any act of intentionality has an aboutness or directed-at-ness quality. Moreover, although rarely analyzed, it is the intentionality of a subject that is directed in a specific manner at the intentional object O. More specifically,

> *Directedness*: Necessarily, if S is in an intentional state P about O, then P is directed in the manner F at O.[104]

Directedness is related to a further thesis:

> *Givenness*: Necessarily, the object O of an intentional act P is given to S of P.

Uwe Meixner argues that *Directedness* entails *Givenness* in the following way.[105]

---

[101] Recent defenders of adverbial non-relational intentionality include Addis, L. (2005). The necessity and nature of mental content. In G. Forrai & G. Kampis (Eds.), *Intentionality: Past and future* (pp. 1–14). Amsterdam: Rodopi B.V. Adverbialism about intentionality has recently been defended as a thesis about phenomenal intentionality, according to which facts about consciousness have an adverbial structure; they are facts about ways of thinking, perceiving. See Banick, K. (2021). How to be an adverbialist about phenomenal intentionality. *Synthese*, *198*, 661–686; Kriegel, U. (2007). Intentional inexistence and phenomenal intentionality. *Philosophical perspectives*, *21*(1), 304–307; and Kriegel, U. (2008). The dispensability of (merely) intentional objects. *Philosophical Studies*, *141*, 79–95.

[102] Plausibly, *Monadic* entails the following. *Monadic Subject Necessity*: for all acts of intentionality P, if P is exemplified by a subject S (as a mental state of S), then, necessarily, P is a mode of S. Our argument, however, does not require this thesis. Although, if *Monadic Subject Necessity* is true, then our argument is strengthened.

[103] Meixner specifically refers to the work of George Bealer and John Searle. Uwe Meixner, "Classical Intentionality," Erkenntnis 65 (2006), 33–34.

[104] In Chapter 9, we will see that *Directedness* does not entail conceptual empiricism.

[105] Meixner, Classical intentionality, pp. 33–34.

> *Givenness Entailment*: If an intentional subject S is directed in the manner F (intrinsic to the intentional experience) at an intentional object O, then, necessarily, O is given in the manner F to S.
>
> *Directedness Entailment*: If an intentional object O is given in the manner F (intrinsic to the intentional experience) to a subject S, then, necessarily, S is directed in the manner F at O.

As we will explain below, *given-to-S-ness* includes an ontology of self-presenting properties with important implications.

5.3.5 Direct Self-Awareness as a Self-Presenting Property

The ontology of the *given-to-S-ness* can be understood as a self-presenting property. As we will show, the given-to-S-ness, understood as a self-presenting property, entails that I am directly aware of myself in intentional acts. To understand our account of dualist seemings, we must grasp the nature of self-presenting properties. Following Roderick Chisholm, we adopt:[106]

> *Self-Presenting*: P is self-presenting $=_{\text{def.}}$ Every property that P entails includes the property of thinking.[107]

Understanding this requires clarifying Chisholm's definitions of property entailment and inclusion.[108] Properties may sustain these relations with each other:

> *Inclusion*: Property P includes property Q $=_{\text{def.}}$ P is necessarily such that whatever exemplifies it exemplifies Q.
>
> *Entailment*: Property P entails property Q $=_{\text{def.}}$ P is necessarily such that, for every *x* and every *y*, if *y* attributes P to *x*, then *y* attributes Q to *x* or, alternatively, believing something to be P includes believing something to be Q.

Inclusion requires that the very same entity that exemplifies a property P must also exemplify the property P includes. Determinates (being red) include their determinables (being colorful). Notice, also, that *Entailment* contains the concept of attribution. In an act of attribution, one ascribes a property to something. A judgment or belief that *x* is F is an attribution of being F to *x*. For example, the property of <believing there are red circles> entails the property of <believing that there are red things>. Self-presenting properties characterize thoughts, sensations, and other states of consciousness. The property of <being-appeared-to-redly> entails the property of <being-appeared-to>, and the latter property includes the property of <thinking>.

---

[106] See Chisholm, R. M. (1977). *Theory of knowledge* (2nd ed., pp. 20–22). Englewood Cliffs, NJ: Prentice-Hall; Chisholm, R. M. (1989). *Theory of knowledge* (3rd ed., pp. 18–25). Englewood Cliffs, NJ: Prentice-Hall; Chisholm, R. M. (1981). *The first person: An essay on reference and intentionality* (pp. 79–83). Minneapolis: University of Minnesota Press.

[107] Chisholm, *Theory of knowledge* (3rd ed., p. 19). While Chisholm proffered slightly different definitions of a self-presenting property, we take this definition as representative of his views.

[108] See Chisholm, R. M. (1989). *On metaphysics* (pp. 143–145). Minneapolis: University of Minnesota Press; and Chisholm, R. M. (1996). *A realist theory of categories* (pp. 11–21). Cambridge: Cambridge University Press.

I can be aware of the external, physical world only by means of my mental states, but I need not be aware of my mental states by means of anything else. In different ways, a self-presenting property presents to or is given to a subject the intentional object of that property and the self-presenting property itself. Such properties present other things to a subject intermediately by means of them, and they present themselves to a subject directly simply in virtue of the fact that he has them.

In normal circumstances, it is by virtue of exemplifying the relevant self-presenting property that a subject comes to have knowledge by acquaintance with the intentional object of the property and with the property itself. Moreover, it is in virtue of exemplifying the relevant self-presenting property that a subject is *prima facie* justified in believing that the intentional object has a particular feature.

For example, it is by way of a sensation of red that one is aware of the surface color of an apple, but one is not aware of the sensation of red by way of another sensation. The red sensation makes the apple's surface present to one by virtue of one having the sensation, but the sensation also presents itself directly to one without another intermediary. A person sees red by means of the sensation of red and is made directly aware of (but does not actually see in the same way he sees the red on the apple's surface) the sensation itself by having that sensation. The person gains knowledge by acquaintance with the property of <being red on the apple's surface> and the property of <being-an-appearing-of-red>. Further, it is on the basis of the latter, viz., that the apple's surface seems red to one, that the person is *prima facie* justified in believing that the apple's surface is red.

5.3.6 Self-Awareness Necessity

When a person exemplifies a self-presenting property, they are modified in some way. We may put this by saying that when a person has a red sensation, the person is in the state of *being-appeared-to-redly*. Suppose the light is such that an orange jar looks red to Smith. If Smith says, "The object is red," his statement is false. If Smith says, "I seem to see something red" or "The jar appears red to me," what he says is true because he is reporting a description of his own sensation. He is not talking about the jar. Smith's statements employ a phenomenological use of "seems" or "appears." When people use "seems" or "appears" phenomenologically, they use them to report their own description of their self-presenting properties, i.e., to report the private, directly accessed mental states going on inside them. Since an instance of such a property is a mode of the self, such a report strictly entails being directly aware of oneself. At the time of instantiating the self-presenting property, the self is modified, and one's self-awareness is precisely an awareness of *one's self thusly modified*. This is what is reported.

If one is appeared-to-redly while looking at a red apple, then this self-presenting property has two intentional objects. First, the *primary intentional object*, the red surface, is the object of focused awareness, in this case, sensory awareness. Such an object requires a distinct mental state to be purposefully directed upon it. However, the self-presenting property also presents itself to the subject as a secondary, *peripheral intentional object*. The subject is aware of the self-presenting property <being-appeared-to-redly> but as the primary intentional object of a distinct mental act. As a secondary, peripheral object, the self-presenting property presents itself to a subject by simply being instantiated by that subject, and it may. However, it need not be the object of a distinct mental act of attending to it.

From this ontology of self-presenting properties, it follows that:

> *Self-Awareness Necessity*: Necessarily, if S exemplifies a self-presenting property *y*, then S is directly aware of S, where S = S.

In the words of Chisholm, "[W]hat one apprehends when one apprehends heat or cold, light or shade, love or hatred, is simply oneself. Whether one knows it or not, one apprehends oneself as being affected or modified."[109]

5.3.7 The Causal-Acquaintance Principle

Why think we are directly acquainted with ourselves? Consider the following principle:

$$(s)(x)(y)(\mathrm{K}_a sx \to \mathrm{K}_a sy),$$

where *s* ranges over knowing subjects, *x* ranges over causal facts, *y* ranges over the associated causal objects that constitute their causal facts; K_a is "has knowledge by acquaintance with." That is,

> *Causal-Acquaintance Principle* (CA): Necessarily, if a subject S has knowledge by acquaintance with a causal fact *x*, then S has knowledge by acquaintance with the object *y* that caused *x*.

As Tim O'Connor illustrates, if S is directly aware of a hammer's causing a nail to move, then S is directly aware of the hammer. CA seems to account for a wide range of cases and is highly justified.

Consider the difference between active and passive thoughts. A *passive thought* happens to me as a patient when I am, say, listening to someone talk. By contrast, an *active thought* is one in which I exercise active power with respect to and entertain freely as an agent. We are quite capable of knowing the difference between active and passive thoughts. Still, as a compatibilist would have it, we do not vouchsafe such knowledge by gaining further knowledge about the causal pedigree of the two types of states. Instead, we are directly acquainted with the difference and can be aware of it by simply attending to the relevant mental states. Take as a causal fact my causing an active thought. When I act as an active agent by freely choosing to think about lunch, for example, instead of being a patient passively triggered to think of lunch by hunger pangs. It satisfies the antecedent of CA. So, I have knowledge by acquaintance with myself.

Lastly, following J. G. Fichte, it seems that at least two undeniable things are impossible if we are not acquainted with—directly aware of—ourselves. First, we couldn't know who or what we direct our intentionality on when we do so on ourselves. Second, there could be no rational justification for directing our intentionality on ourselves, as we would have no reason to believe there is anything to direct our intentionality.[110]

---

[109] Chisholm, On the observability of the self, p. 19.

[110] Zagzebski, L. T. (2021). *The two greatest ideas: How our grasp of the universe and our minds changed everything* (p. 145). Princeton, NJ: Princeton University Press. Zagzebski refers to: Henrich, D. (1982). Fichte's original insight. In D. E. Christensen (Ed.), *Contemporary German philosophy* (Vol. 1, pp. 15–53). University Park: Pennsylvania State University Press. Neuhouser, F. (1990). *Fichte's theory of subjectivity*. Cambridge: Cambridge University Press.

5.3.8 Monadic Intentionality, Direct Self-Awareness, and Dualist Seemings

In this Chisholmian view, by way of a self-presenting property, a subject can take other entities as objects of intentionality, including as objects of propositional knowledge. In different ways, self-presenting properties present their intentional objects and themselves to first-person knowing subjects. Uwe Meixner explains the vital role of the intentionality of self-awareness and the formation of beliefs about oneself.

We certainly do not usually attribute beliefs to ourselves in order to explain our behavior to ourselves or to others. We usually attribute beliefs to ourselves because we have certain *intentional experiences: belief-experiences* (usually not in isolation, but interwoven with other intentional experiences of ours, for example, with other beliefs and, very often, with perceptions). These belief-experiences are certainly rather far from being theoretical constructs—and so are the *beliefs* whose vehicles the belief-experiences are.[111]

Applying *Monadic Intentionality*, when one is aware of oneself (as opposed to a mental state one has), one simply directs one's intentionality onto oneself, such that the subject and object of awareness stand in the identity relation to each other. In an act of self-directed intentionality, one is directly acquainted with oneself and certain self-presenting properties of oneself, including features revealed in acts of dualist seemings. In acts of self-directed intentionality, one has direct self-awareness—knowledge by acquaintance with oneself—that reveal oneself as an immaterial simple substance.

According to Brentano, not only can we can know ourselves but ourselves as a substance. Only one thing can immediately perceive impressions with inner evidence. Even in the act of assuming that we know something about ourselves we know ourselves as substances.[112] *Monadic Intentionality* makes sense of this. In acts of self-directed intentionality, we are aware of ourselves as modified (in our accidents, e.g.) in specific ways. But to know that one is modified is to know that oneself is a substance and modes of oneself.[113] This explains how dualist seemings result from knowledge by acquaintance with oneself and those features of oneself that reveal the truth of SD and the falsity of physicalism and materialism.

5.3.9 Satisfying the Explanatory Conditions of Dualist Seemings

DSA easily accounts for the three explanatory conditions mentioned above (II.2). DSA satisfies *Diversity* by not reducing dualist seemings to one seeming described in different ways and recognizing the full data set (i)–(xviii) and in an explanatorily significant way.

Ubiquity is, on DSA, explained by the fact that knowledge of the self is present in normal states of phenomenal consciousness. Dualist seemings are pervasive because they are pre-theoretical, which is what we should expect if we have direct self-awareness.

---

[111] Meixner, Classical intentionality, p. 29.

[112] See, e.g., Rancurello, A. C. (1968). *A study of Franz Brentano: His psychological standpoint and his significance in the history of psychology*. New York: Academic Press.

[113] For a fascinating argument that the nature of intentionality and mental properties logically require a subject (understood as a substance and not a bundle of properties) that is modified and thus fundamentally conscious, see Chisholm, R. M. (1991). The bearers of psychological properties. *Aaiucov*, *3*, 7–14.

DSA explains *Strength* by direct self-awareness with one's self. This is so because they express how things phenomenologically seem or appear to us. Not only is one directly aware of one's mental states, but also of one's self. We have these intuitions by virtue of our direct awareness of ourselves as simple, spiritual substances. As we've explained above, knowledge by acquaintance is epistemically stronger than propositional knowledge. Dualist seemings are so strong because they are self-presenting or given in one's direct self-awareness. We know—*via dualist seemings*—certain things about our mental states and about ourselves better than we know the premises of arguments for naturalist views. Hence, dualist seemings provide us with Moorean facts, and when they do, the dualist needn't revise her ontology in the face of arguments for physicalism.

5.3.10 Explanatory Unity

In addition to explaining *Diversity*, *Ubiquity*, and *Strength*, our account of dualist seemings provides explanatorily unity. It often happens in science that a range of apparently unrelated data can be explanatorily unified if a theoretical entity is postulated as that which is causally responsible for that range. More specifically,

> *Explanatory Unity*: A theory *x* has unifying power with respect to a range of phenomena p_1-p_n, *iff*, *x* sufficiently explains p_1-p_n by grounding facts about p_1-p_n in facts about a single entity, *e*.

The postulation of electrons, for example, unified a wide range of phenomena by depicting them as effects of the electron's causal powers. Explanatory unifying power solves the problem of explaining why a range of phenomena are as they are without leaving them as a rather odd "group" of brute facts by explaining that facts about the phenomena are all grounded in facts about the same entity. In solving these problems, explanatory unity satisfies an essential desideratum for both ontological and scientific explanatory theories.[114]

Similarly, a range of apparently unrelated items of knowledge can be unified by one's knowledge by acquaintance with some relevant object. For example, I know many things about a certain spatial region R in the classroom in which I lecture. I know that everyone walks around R and not through it. R is rectangular, four feet tall, contains something that is darker than yellow, and that if a book is placed near the top of R, it will hover in stable suspension off the floor. But rather than being a set of isolated pieces of knowledge, they have a unifying pattern. I know each of them in virtue of knowledge by acquaintance of the podium that overlaps R. It is because of that acquaintance that I know these features of R.

Our DSA account unifies the diverse dualist seemings (i)–(xviii) by grounding facts about (i)–(xviii) in facts about one's direct self-awareness. Moreover, the lack of explanatory unity on non-SD accounts is evident from the fact that rival solutions to the meta-problem of consciousness fail to cover all dualist seemings.[115] Chalmers points out, among other examples, that many explanations of dualist seemings do not apply to zombie intuitions. The same is certainly not true of explicit SD seemings (xi)–(xviii).

---

[114] It is important to note that explanatory unity is not the same as explanatory simplicity, which we argued in Chapter 2 should be understood epistemically. Explanatory unity is epistemically prior to simplicity because its application to a range of phenomena ontologically unifies them in such a way that the phenomena are no longer a catalogue of disparate facts/entities. The gain of explanatory unity is not primarily or merely related to explanatory simplicity. We will explore this in more detail in Chapter 10.

[115] See Chalmers, D. J. (2020). How can we solve the meta-problem of consciousness? *Journal of Consciousness Studies*, *27*(5–6), 201–226.

5.4 A DIRECT REFLECTIVE SELF-AWARENESS ARGUMENT FOR SUBSTANCE DUALISM

We've offered an SD solution to the hard meta-problem of consciousness; an explanatorily powerful account of dualist seemings, grounded in knowledge by acquaintance with one's self in a phenomenologically faithful and uncontrived way. There is, however, a further and significant implication of our account. We will now provide a robust analytic phenomenological argument that in addition to providing *prima facie* (initial appearance) justification for SD, dualist seemings provide *secunda facie* (upon further inspection or second appearance) justification or defeasible knowledge that SD is true. This account has the advantage of remaining phenomenologically faithful while avoiding the error of reducing dualist seemings to mere dispositions.

5.4.1 Husserlian Fulfillment Structures and Self-Knowledge

Acquaintance is similar to Husserl's notion of experience with fulfilled intention. Contrary to Bertrand Russell, Husserl held that we could have authentic representation with fulfilled intentions with reference to "sense-objects" (including the self, internal experience, external objects, and their perceptual qualities) rather than sense-data.[116] For Husserl, knowledge involves the recognition of an identity between an object adequately intuited and the object as merely thought. The conscious awareness of this identity is called "fulfillment." Fulfillment, explains Dallas Willard, is the "union of the conceptualizing act with the object, on the basis of a corresponding intuition of that object together with a recognition of the identity of the object of the concept and of the perception."[117] So, fulfillment is knowledge.

5.4.2 Knowledge and Intentionality

Fulfillment is usually the final mental act of a diachronic series of relevantly ordered mental acts that lead a subject from an empty intention to an adequately filled intention in which the object known is directly, intuitively present as it really is. Such a series is called a "fulfillment structure." In such cases, one forms a concept of something (e.g., that a specific book is waiting to be picked up in the university bookstore) and goes through a series of experiences (walking to the store, entering the store, walking closer and closer to the relevant section of the store, seeing the book from a distance) that terminate in the experience of comparing one's initial concept of the thing to the thing itself to see if there is a match.

---

[116] We owe this connection to Dallas Willard. See Willard, D. (1984). *Logic and the objectivity of knowledge: A study in Husserl's early philosophy* (pp. 91–92, and 232–234). Athens, OH: Ohio University Press.

[117] Willard, D. (1995). Knowledge. In B. Smith & D. W. Smith (Eds.), *The Cambridge companion to Husserl* (p. 152). Cambridge: Cambridge University Press.

For Husserl, *de re* knowledge obtains when a certain type of immediate intuition/presentation of (direct acquaintance with) the thing itself. Indirect *de re* knowledge involves inference(s) that trace back to immediate or *originary* presentations. *De dicto* knowledge is the conscious possession of truth/adequation on an appropriate immediate or mediate basis in thought and experience.[118] *Immediately based knowledge* is the possession of truth by virtue of direct encounter with the facts/things themselves, as they are given in acts of consciousness.

What is important about fulfillment structures is that from a first-person perspective, one can grasp one's initial (vague) concept (what Husserl called an *empty intention*), understand (perhaps somewhat vaguely) what its intentional object is, infer a series of experiential steps that would help to verify or disconfirm the initial concept, and eventually compare the initial concept with the intentional object. Fulfillment structures and the cognitive success to which they (often) lead require that whatever account of intentionality one gives, it must be such that

(a) The essence of conceptual *meanings* can be grasped from the first-person perspective.
(b) The *intentionality* of conceptual meanings and the nature of their objects can be grasped from a first-person perspective.
(c) The *enduring I* that remains identical through the series of mental events can experience the initial somewhat empty concept become fuller and fuller until the thing itself is present in knowledge by acquaintance.

5.4.3 Fulfilled SD Seemings

Now, when I reflect upon my experience I conceptualize the experience and its properties appropriately. So, for instance, I can immediately verify (as opposed to infer) propositions like "My experience is of a table," "My experience is inadequate to its object," "My experience is intuitive rather than empty," "My experience posits the table as an actual (as opposed to fictitious) object," and so on. I can also fulfill the thought that "My experience of the table is not directed upon an object O* which I take to be of the table." My experience is direct, intuitive, positing, inadequate, or incomplete, and originary.

Applied to self-knowledge, in reflection, I apprehend myself *intuitively* and *originarily*. My self and its states or qualities are given to me in the strictest and most adequate possible way; they are known by acquaintance. When I think of or reflect on those states, I undergo a case of fulfillment. I come to know that I am a simple, immaterial substance, among many other things.

That *dualist seemings* arise in virtue of direct self-awareness explains both the Strength and Ubiquity of dualist seemings in acts of inner-consciousness or reflection. Husserl's account of knowledge, or fulfillment, explains why the *Strength* of dualist seemings would or would not naturally attach itself to an act with fulfilled or unfulfilled content. Hopp explains,

> When S merely thinks that P [an unfulfilled content], the object is merely intended and not given. It itself does not show up, and it is impossible to check whether the content is true by measuring it against "the things themselves."

---

[118] Willard, D. (2000). Knowledge and naturalism. In W. L. Craig and J. P. Moreland (Eds.), *Naturalism: A critical analysis* (pp. 24–48). London: Routledge, especially 31–34.

> The emptiness of the intention is a positive phenomenological feature of the act. The *absence* of the object, and the corresponding lack of any *real* relation to it, is an experienced lack or shortcoming. This *explains why* P will not seem true to a rational person when it is merely entertained in the absence of any other supporting considerations.[119]

The experience of an empty intention is so phenomenologically striking that it drives our natural desire to seek knowledge, to seek fulfillment.[120] This explains the Ubiquity of dualist seemings on our DSA model. Were one not aware of oneself or not aware of oneself as simple immaterial substances, then one would not experience the fulfillment of their intention; the intention would be empty. But that isn't what people experience. Instead, we experience, in reflection, the fulfilment of our dualist seemings.[121] In reflection, we know by direct acquaintance of our self that SD is true.

5.5 OBJECTIONS TO SELF-AWARENESS ARGUMENTS

We dealt with the prominent objections to introspection in the previous chapter. This section addresses the most likely objections to our argument from direct self-awareness.

5.5.1 Hume's Phenomenological Report

Aristotle famously held that a person must, while perceiving anything, also perceive their own existence (*De Sensu* 7.448a). However, the most obvious objection to DSA is to echo the phenomenological report of David Hume:

> For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe anything but the perception. When my perceptions are remov'd for any time, as by sound sleep, so long am I insensible of myself, and may truly be said not to exist.[122]

Although contrary to the commonsense consensus, Hume's phenomenological report has its defenders. However, Hume's view faces various problems.

---

[119] Hopp, *Phenomenal conservatism and the principle of all principles*, pp. 194–195. Hopp references the following. Willard, D. (1984). *Logic and the objectivity of knowledge* (227). Athens, OH: Ohio University Press; and Bernet, R. (2003). Desiring to know through intuition. *Husserl Studies, 19*, 153–166. This is one way Husserl's view differs from, and is superior to, phenomenal conservatism, a point that Hopp develops in great detail.

[120] Husserl, *Logical investigations*, p. 722. Credit to Willard for this reference.

[121] Walter Hopp develops this point in detail. See Hopp, W. (2015). Phenomenal conservatism and the principle of all principles. In D. O. Dahlstrom, A. Elpidorou, & W. Hopp (Eds.), *Philosophy of mind and phenomenology: Conceptual and empirical approaches* (pp. 180–202). New York, NY: Routledge.

[122] Hume, D. (2000). *A treatise of human nature* (D. F. Norton & M. J. Norton, Ed.). Oxford: Oxford University Press. (First published, 1739), Sect. 6.3, SB252.

Here is one way to state the most difficult problem. Chisholm poses the following question.

> How can he say that he doesn't find himself—if he is correct in saying that he finds himself to be stumbling and, more fully, that he finds himself to be stumbling on certain things and not to be stumbling on certain other things?[123]

Suppose Hume denies the existence or identity of the self by offering a phenomenological report of failing to observe a self (the experience of an empty intention) and observing only perceptions. In that case, there is literally nothing to which this phenomenological report belongs. The only way to keep the Humean argument is to understand the self as a bundle of perceptions that "form a whole only by being connected together."[124] But then we have no Humean rejection of the self.

5.5.2 Commonsense Materialism

John Searle describes his version of materialism as "a kind of commonsense answer to the question, given what we know about how the world works." Still, in the following sentence Searle admits that his materialism is "very much a minority view."[125] Stewart Goetz and Charles Taliaferro list Gilbert Ryle, Anthony Flew, and P. M. S. Hacker among those who deny that SD is a commonsense view.[126] Animalists often talk this way.[127] Trenton Merricks argues from the commonsense belief that he kisses his wife. Yet, if SD is true, he has never kissed his wife. Only his body can kiss his wife's body.[128] But this violates common sense and reveals the speciousness of materialism.

That materialism is more in line with common sense than SD, argues Charles Taliaferro, is based on a misconception of SD, consciousness, and the explanatory nature of physical sciences.[129] As mentioned in Chapter 1, Ryle's "ghost in the machine" mischaracterization of SD ignores the embodiment views of SD defenders. We will have much more to say about this in Chapters 10 and 11.

The simple fact is that it does not seem to most folk that they are macro-level objects. On the contrary, it seems to most people from the first-person perspective that we are mental subjects who fail to be aware of exemplifying any physical properties. The issue then becomes whether there is any good reason to think we are physical objects, though we are not aware of being such.

---

[123] Chisholm, R. M. (1969). On the observability of the self. *Philosophy and Phenomenological Research*, *30*, 10.

[124] Hume (2000), 400.

[125] Searle, J. (1984). *Minds, brains, and science* (p. 28). Cambridge, MA: Harvard University Press.

[126] See Goetz, S., & Taliaferro, C. (2011). *A brief history of the soul* (pp. 182–186). Malden, MA: Blackwell.

[127] For example, animalist Peter van Inwagen claims that when he enters most deeply into himself, he discovers that he is an animal. See van Inwagen, P. (1998). Modal epistemology. *Philosophical Studies*, *92*, 67–84.

[128] Goetz and Taliaferro, p. 186.

[129] Taliaferro, C. (2014). The promise and sensibility of integrative dualism. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 191–203). New York, NY: Routledge.

5.5.3 Linguistic and Theoretical Dependence

Others have argued that our most basic intuitions about ourselves are dependent on the language we speak, and this language is derived from past and present theories about the way things are. This objection is mistaken for at least two reasons. First, as the empirical and phenomenological data shows, dualist seemings are not primarily philosophical or theoretical; they are commonsensical seemings.

Moreover, this account of the origin of dualist intuitions has it backward. People did not create the idea of the soul as a part of their theorizing about the world; their theorizing about the world employed and extended what they already pretheoretically (and prelinguistically) knew. Commonsense SD expresses what Jesse Bering calls a "folk psychology of souls," not mythology.[130] Ilkka Pyysiäinen observes,

> All peoples have beliefs about various types of souls that are responsible for the liveliness of the body as well as for various cognitive-emotional functions ... Such conceptions thus cannot be analyzed merely at the level of their linguistic expressions and are not spread only through cultural contacts.[131]

Consider the idea that people theorized another life as a place where punishments and rewards are justly distributed. Where did these people come up with the idea that they were the sorts of things that could survive in an afterlife? Where did the very idea of an afterlife come from, and why was it so widely believed to be disembodied? As Pyysiäinen explains the anthropological data of early cultures,

> Although humans may have lacked doctrines and narratives about souls, they must have had intuitive ideas about minds ... Otherwise, even a rudimentary form of religion would not have been possible.[132]

After a detailed analysis, Raymond Martin and John Barresi conclude that commonsense SD "did not even begin with the Greeks or Hebrews, or even the Egyptians. Rather, it began much earlier, perhaps with the Neanderthals ..."[133] In our view, people's notion of an afterlife was already justified as a possibility because of the DSA account of dualist seemings we've considered, and these dualist seemings were subsequently pressed into service, not the other way around.

5.5.4 Absent Dualist Seemings

Some may reply that there are introspectively attentive people who do not have these seemings. But this is surely a strained reply. The overwhelming majority of people now and throughout history have held the intuitions listed above. And at the beginning of this chapter, we noted that at least pre-philosophically, many, perhaps most physicalists have these intuitions. The usual reason for rejecting these intuitions is a question-begging, prior commitment to physicalism which sets these intuitions and their empirical support aside (usually eliminated or reduced).

---

[130] E.g., Bering, J. M. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, *29*(5), 453–462; and Bering, J. M. (2002). Intuitive conceptions of dead agents' minds: The natural foundations of afterlife beliefs as phenomenological boundary. *Journal of Cognition and Culture*, *2*(4), 263–308.

[131] Pyysiäinen, I. (2009). *Supernatural agents: Why we believe in souls, gods, and buddhas* (p. 93). Oxford: Oxford University Press.

[132] Ibid., p. 68.

[133] Martin & Barresi, *The rise and fall of soul and self*, p. 290.

5.5.5 Religious Etiology

Others may retort that SD is the commonsense view due to religious indoctrination. This retort carries little weight. First, this explanation is empirically undermined. Barret's research, for example, has demonstrated widespread SD beliefs even within self-professing atheistic communities.[134] This same point is true of little children prior to any religious training. We need an explanation for why this is the case. Self-awareness provides such an explanation. Second, this retort gets the explanatory relation backward. Religious SD teachings are not only acceptable to people worldwide because they capture what people already know about themselves due to self-awareness; it is not the other way around.

5.5.6 Neuroscientific Defeaters

Some might sight studies where high doses of a hallucinogenic produce an experience of ego dissolution or a lack of self-awareness. Even on the assumption that these cases show states of consciousness without self-awareness, these studies also suggest that ordinary conscious experience involves a minimal kind of self-awareness, which is what appears to fade away.[135] Our argument does not require the thesis that all states of consciousness involve an awareness of the self, but only that dualist seemings are explained by direct self-awareness revealing the truth of SD. Moreover, if there is no self we are aware of, we would expect people to report no self, especially those in neuroatypical conditions. However, various empirical studies reveal the opposite, even in cases of thought-insertion in schizophrenia.[136]

5.5.7 Self-Awareness Failure

Where we appeal to positive, direct awareness of the self, critics will appeal to a failure to be aware of substantive mereological complexity (a failure to be directly aware of one's separable parts). Moreover, they will likely point out that this failure better explains why so many philosophers (e.g., materialists like Jackson and Nagel) who accept most or all dualist seemings (i)–(xvii) are not dualists. Our claim to be aware of the substantive simple self fails to provide any explanation for why these scholars are not dualists. Indeed, if we are correct, these people should be SDs.

---

[134] Barret, *Born believers*, p. 212.

[135] E.g., Millière, R. (2017). Looking for the self: Phenomenology, neurophysiology and philosophical significance of drug-induced ego dissolution. *Frontiers in Human Neuroscience*, *11*(245), 1–22. doi: 10.3389/fnhum.2017.00245.

[136] E.g., Duncan, M. (2019). The self shows up in experience. *Review of Philosophy and Psychology*, *10*(2), 299–318.

We have three things to say in response. First, these physicalists lack the second-order belief that they have the first-order awareness of themselves as simple immaterial substances because they are looking for the wrong sort of experience constitutive of that first-order awareness. Phenomenologically, if a philosopher claims to be directly aware of some entity, another philosopher can always claim not to have that awareness. When G. E. Moore claimed to have a direct intuition of goodness, his rivals simply denied they had the same intuition. A defender of Moore could respond in this way: Goodness is a second-order property like being colored or being shaped, not a first-order property like pleasure, being red, or being triangular. Now intuitions of these sorts of first-order properties have a certain texture or vivacity that is absent in the case of intuitions of the corresponding second-order properties. Those who failed to have the relevant intuition of goodness were looking for the intuitive texture appropriate to the intuition of a first-order property. Of course they never found it. Unfortunately, they were looking for the wrong sort of phenomenology. Once an intuition of goodness is compared to the intuition of other second-order properties, it becomes more plausible to think that the relevant intuition is real.

Something like this is going on with respect to the direct awareness of the content of propositional attitudes, for example, thoughts and beliefs. Just as we can grasp by direct self-awareness the nature of a pain (and other states of phenomenal consciousness), we can directly grasp the introspectively available nature of the conceptual and propositional contents that constitute our thoughts, beliefs, and so forth.

Jaegwon Kim disputes this claim, arguing that a kind of belief content, for example, <that George Washington was the first president of the United States>, does not have a uniform, qualitatively introspectable character present in all instances of that kind.[137] However, Kim is guilty of a Humean vivacity test for a phenomenal quale (for example, a conceptual content), according to which one has such a quale only if it is a vivid sensation like a pain or an image of George Washington. But in this case, Kim's belief about this matter is based on looking for the wrong sort of first-order awareness. He mistakenly thinks that if a direct awareness of the content of a thought fails to have the sort of vividness of a sensation such as a pain, then there is no such awareness. However, that the introspective texture of a thought's content is not as vivid as that of a pain does not entail that we do not have the former. Those who think otherwise are looking for the wrong sort of awareness.

Now, in our view, something like this is going on with respect to the positive first-order awareness of the self. People who reject such awareness on the grounds that they simply do not find that they have it are looking for something like a sensation of pain or some other vivid mental state. But an awareness of the self is not like that, which is one reason why people mistakenly believe they do not have the awareness.

Our second response rests on the notion that while intuitions, construed as a way things seem phenomenologically to a subject, are harder to change than beliefs, beliefs can cause someone to reject an intuition. This can occur when an intuition whose corresponding belief is *prima facie* justified to a subject on the basis of that seeming is overridden by a weightier belief. Thus, while dualist seemings (*i*)–(*xvii*)—and SD itself—receive *prima facie* justification from the relevant more basic seemings produced by or a constituent of the phenomenal state of direct self-awareness of one's self as a simple immaterial substance, physicalists do not take themselves to have this more basic phenomenal state because of what they think is an overriding defeater. What is that overriding defeater? In the context of debates about property dualism regarding qualia, Kim says the following:

---

[137] Kim, J. (2011). *Philosophy of mind* (3rd ed., pp. 275–276). Boulder, CO: Westview. Paul Churchland goes so far as to argue this same point with respect to mental states like pain that are clearly categorized as states of phenomenal consciousness. See also Churchland, P. (2013). *Matter and consciousness* (3rd ed., pp. 89–90). Cambridge, MA: MIT Press. Unfortunately, Churchland fails to see that painfulness is a second-order property, not a first-order property, and as such, it characterizes various kinds of pain that differ phenomenologically with respect to their species and not their genus.

> The [property-dualist] case against qualia supervenience therefore is not conclusive, though it is quite substantial. Are there, then, considerations in favor of qualia supervenience? It would seem that the only positive considerations are broad metaphysical ones that might very well be accused of begging the question.[138]

Kim goes on to say that these broad metaphysical considerations amount to the assumption that physicalism must be true. He also correctly claims that *the corresponding dualist intuitions are not based on a prior commitment to property dualism but, rather, provide justification for property dualism*. We believe that the same question-begging prior commitment to physicalism funds the rejection of direct awareness of the self as a simple immaterial substance.

To avoid the charge of begging the question, the physicalist must find independent reasons for physicalism sufficient to override what we seem to be aware of in first-person introspection. In our view, the philosophical arguments for physicalism are surprisingly weak. And in any case, physicalists usually rest their case on the supposed findings of science. We will argue at length in Chapter 10 that this is mistaken.

5.5.8 The Dual-Process Defeater

One might argue that dualist seemings have a more plausible neuroscientific explanation, one that defeats DSA. The only developed objection of this kind is aimed at explanatory gap seemings. We will focus on this proposed defeater.

There are several empirical studies related to the explanatory gap from work in the theory of mind,[139] consciousness studies,[140] and mental state attribution,[141] including a recent meta-analysis of these studies taken together.[142] We are interested in analyzing a popular account advanced by cognitive scientists:

> On our view, the explanatory gap works [because] people's natural inclination to judge that broadly physicalist accounts of consciousness "leave something out" depends on a cognitive architecture involving two distinct processes. In typical

---

[138] Kim, *Philosophy of mind* (2nd ed., p. 233).

[139] Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press; and Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and "theory of mind." *Mind and Language*, *14*, 131–153.

[140] Block, Troubles with functionalism. and Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, *18*, 227–247.

[141] Inagaki, K., & Hatano, G. (1991). Constrained person analogy in young children's biological inference. *Cognitive Development*, *6*, 219–231; and Johnson, S. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences*, *4*, 22–28.

[142] E.g., Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and Cognitive Sciences*, *7*, 67–83.

> cases of consciousness attribution, the processes produce harmonious outputs and lead to unsurprising attributions. However, in the case of the explanatory gap, we claim, one of the relevant cognitive processes fails to produce any output, thus leading to the disharmonious sense that the neural description is fundamentally incomplete as an explanation of consciousness.[143]

If successful, this reply would undermine the evidence of explanatory gap seemings.

This defeater assumes a structural model of cognition according to which information processing involves two process types, what Daniel Kahneman calls System 1 and System 2 processes.[144]

> *System 1 Processes*: Quick, automatic, unconscious, associative, heuristic-based, computationally simple, evolutionarily old, domain-specific, and non-inferential.
>
> *System 2 Processes*: Relatively slow, controlled, introspectively accessible, rule-based, analytic, computationally demanding, inferential, domain-general, and voluntary.

System 1 processes are less prone to error, although both systems are reliable in their own respects.[145] Moreover, these systems often operate over the same domain.

According to Fiala et al., attributions of consciousness happen through one of the following two cognitive pathways with System 1 or System 2 characteristics.[146]

> *High-Road Processes*: The judgment that *x* is conscious by a system of rational deliberation, theory application, or conscious reasoning.
>
> *Low-Road Processes*: The judgment that *x* is conscious by a system that is fast, domain-specific, and automatic.

Research suggests that we typically attribute consciousness, via Low-Road cognitive processes, by identifying three surface-level features: appearing to have eyes, appearing to behave in a contingently interactive manner, and displaying distinctive (non-inertial) motion trajectories. This account is referred to as AGENCY. The presence of an object's AGENCY features elicits a "gut-feeling" that the object is conscious. Objects lacking AGENCY characteristics (e.g., the brain or neural events) do not trigger consciousness attribution.[147]

---

[143] Fiala, B., et al. (2011). On the psychological origins of dualism: Dual-process cognition and the explanation gap. In E. Slingerland & M. Collard (Eds.), *Creating consilience: Integrating the sciences and the humanities* (p. 99). Oxford: Oxford University Press.

[144] Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, *93*, 1449–1475.

[145] Ellen Peters, for example, has shown that people who possess the capacity and disposition to use System 2 reasoning quite well do not use System 1 processes less when outperforming others. In fact, they use more of both systems. Their System 2 response triggers the use of System 1 reasoning. See, e.g., Peters, E., et al. (2006). Numeracy and decision making. *Psychological Science*, *17*, 407–413; and Peters, E., Slovic, P., & Gregory, R. (2003). The role of affect in the WTA/WTP disparity. *Journal of Behavioral Decision Making*, *16*, 309–330.

[146] Fiala, et al., p. 91.

[147] Ibid., p. 92.

Frequently, Low-Road and High-Road cognitive processes issue discordant consciousness attributions.[148] Individuals accurately attribute consciousness to other persons via Low-Road and High-Road processes functioning in harmony. However, this harmony is missing in explanatory gap cases.[149] Through Low-Road processes, we consider the mass of gray matter lacking AGENCY features and fail to see it as conscious. However, when physicalists conclude, from internalized theoretical beliefs, that physical states are identical to or give rise to consciousness, they do so according to High-Road processes.

According to Fiala et al., explanatory gap intuitions obtain because the neurological evidence for a brain being conscious (e.g., the presence of neurons firing between 40 Hz and 60 Hz), lacks AGENCY qualities. The Low-Road process is not triggered. However, if we possessed such a Low-Road process, and it delivered the right result, then physicalism wouldn't seem false.[150] This gives us the following account:

> *The Dual-Process Account* (DPA): Explanatory gap intuitions are instances of an absent Low-Level cognitive process for attributing consciousness to entities without AGENCY features.

As such, DPA provides undercutting evidence for explanatory gap seemings.

At this point, one might reject DPA by rejecting the dual-processing model itself. The empirical literature for such an argument is readily available.[151] This is a plausible reply. Here are a few more replies to DPA.

5.5.8.1 Dualist Seemings and Moorean Facts

Recall the Strength feature of dualist seemings. As we explained, dualist seemings are so strong that they don't yield to a revisionary physicalist ontology. Yet, the DPA objection demands otherwise. However, revising one's ontology may not be rational if doing so requires that one abandons the obviously true or a Moorean fact.

The DPA account depends on the evidential value of one's first-person direct self-awareness. It is the starting point for any plausible account of consciousness attribution. Without it, we cannot specify what we attribute to others. So, our evidence for DPA depends on our evidence for AGENCY, which depends on our evidence for successfully attributing consciousness to an object without AGENCY traits, our self. This means that any plausible AGENCY account cannot undermine or defeat our seemings about an object being conscious due to a lack of AGENCY features. All that AGENCY tells us is that we regularly attribute consciousness to objects with AGENCY features. DPA then posits that we fail to attribute consciousness because we lack the cognitive process to attribute consciousness to an object without AGENCY features. But this is clearly false when one attributes consciousness to oneself.

---

[148] For example, through High-Road processes, one might reject arguments for other minds, and at the same time, through Low-Level processes, have reactions to people around them as consciousness agents (Fiala, et al., p. 96). In one study, individuals were triggered through Low-Level processes to attribute consciousness to insects. The same individuals, through High-Level processes, based on facts about insect neural systems, rejected the claim that insects can feel pain (p. 94). Many think these conflicting results reveal a cognitive defect in consciousness attribution.

[149] Fiala, et al., p. 96.

[150] Ibid., pp. 97–99.

[151] E.g., Samuels, R. (2009). The magical number two, plus or minus: Dual process theory as a theory of cognitive kinds. In K. Frankish & J. St. B. T. Evans (Eds.), *In two minds: Dual processes and beyond* (pp. 129–146). Oxford: Oxford University Press.

One might object that this reply fails as it has to do with why we do attribute consciousness, while DPA is an account of why we don't attribute consciousness. First, we are not convinced that this is the case. As we've argued, DPA rests on consciousness attribution to our self. For those unconvinced, we offer two more replies.

5.5.8.2 A Dualist Rival to DPA

Here is a second argument. The dualist needn't accept DPA by providing an equally plausible account. Here is a possibility. In cases of explanatory gap intuitions, individuals host at least two intentional objects:

> *Third-p*: From the *third-person perspective*, individuals are aware of the composite nature of a physical thing or system (e.g., a brain and central nervous system).
>
> *First-p*: From the *first-person perspective*, individuals are aware of themselves as unified, as simple (or one of the other features mentioned below).

Notice two things. First, *third-p* and *first-p* have contradictory qualities regarding their intentional structure. *First-p* is an instance of self-awareness, while *third-p* is not. All leading accounts of self-awareness acknowledge that self-awareness is essentially indexical to the subject.[152] Consequently, in explanatory gap cases, individuals experience a shift in awareness from the first-person self-awareness to the third-person perspective as they move between *first-p* and *third-p*. This shift explains the disjointed phenomenological nature of explanatory gap cases. Moreover, this awareness is something we rightly understand as unique to our conscious mental life. We do not stand in this kind of relation to things other than our mental states.

Here is the second problem. The intentional objects of *first-p* and *third-p* present us with contradictory qualities. I am aware of myself as simple.[153] However, the brain, indeed any physical system, is perceived as an object composed of separable parts. What seems obvious to so many is that they are a single, unified self. What is also apparent is that physical objects like the brain are not simple. Leibniz's Law does the rest.

We can offer the same argument with additional mental states features. As mentioned in Chapter 4, I know by introspection that my mental states stand to me as inner (not spatially inside, but properties they exemplify); are private (known by the subject in a way not available to anyone else); are directly given to me (self-presenting); intentional; and necessarily, states of mine.[154] Mental states can be introspectively known and, thus, stand in a direct access relation to the subject.[155] Mental states are inner—exemplified within the subject—and are necessarily states—modes—of the subject.[156] Mental states are intentional; they are about something. Physical things, like brain states, lack these features.[157] These facts become strongly apparent in dualist seemings, including explanatory gap seemings.

---

[152] Gertler, *Self-knowledge*, p. 217.

[153] For a defense of this, see Moreland, J. P. (2011). Substance dualism and the argument from self-awareness. *Philosophia Christi*, *13*(1), 21–34.

[154] John Searle makes a similar argument in Searle, J. R. (2004). *Mind: A brief introduction* (p. 68). Oxford: Oxford University Press.

[155] See Moreland, J. P. (2003). The knowledge argument revisited. *International Philosophical Quarterly*, *43*(2)(170), 219–238.

[156] See Lowe, E. J. (1996). *Subjects of experiences*. New York, NY: Cambridge University Press.

The defender of subject-simplicity can interpret DPA data as cases of attributing consciousness, not to the physical body with the requisite AGENCY features, but to the person who is embodied. Individuals attribute consciousness *to persons* with AGENCY features, not to the body that evidences those features.[158] The empirical data is underdetermined.

5.5.8.3 A Generality Problem for DPA

Here is a third problem: DPA suffers from a problem like the generality problem for process reliabilism about justification or knowledge.[159] DPA is so general that it undermines uncontroversial common sense and otherwise evidentially supported beliefs.[160] And there appears to be no non-question-begging way for DPA to avoid this problem.

DPA gives us a principle by which we should disregard our intuition that something is not conscious. If DPA undermines an intuition that *x* is not conscious, then that intuition does not justify thinking that *x* is not conscious. Although intended to undermine explanatory gap intuitions, DPA extends to various cases, some of which are absurd. Other material objects, for example, thermostats and clocks, lack AGENCY features.[161] According to DPA, we fail to attribute consciousness to thermostats in virtue of the same cognitive account for why we fail to do so for brain states. Hence, DPA requires us to disregard thinking that a thermostat or clock is conscious until we gain evidence, in addition to it seeming that it is not conscious.

So, if DPA undermines our judgment that brain states cannot be conscious, it also undermines our hesitancy in attributing consciousness to clocks and thermostats. This seems wrong! Surely, I am justified in believing that watches and thermostats are not conscious in virtue of it seeming so to me. That DPA undermines these commonsense beliefs means we must reject DPA.

Perhaps the physicalist can avoid this problem by revising DPA. Any revision must limit DPA to a very high degree of specificity by providing criteria that rule in the right material objects and rule out the wrong material objects. The non-reductive physicalist may have such a criterion in emergentism. Broadly speaking, ontological emergence (emergence, for short) is the thesis that when the parts of a system are arranged in the right way, their collective activity brings something fundamentally novel into existence. According to emergentism, objects become candidates for consciousness when they reach a particular level of complexity, which is at least in some cases, achieved in biological systems with a properly functioning brain and brain stem. The emergentist criteria can amend DPA to include only material objects with a certain degree and type of complexity. Clearly, the emergentist criteria rules out thermostats and watches while ruling in organisms with a functioning brain and brain stem.

---

[157] See, e.g., Bealer, G. (1993). Materialism and the logical structure of intentionality. In Robinson, H. (Ed.), *Objections to physicalism* (pp. 101–126). Oxford: Clarendon; and Lowe, E. J. (1995). There are no easy problems of consciousness. *Journal of Consciousness Studies*, 2(3), 206–271.

[158] This is analogous to saying the expression “I am 5 foot 9” merely commits me to being intimately connected to some 5-foot-9 object, not that I am numerically identical to some 5-foot-9 object. Physicalist Eric Olson admits that this dualist reply, “does not conflict with anything that no sane person could deny.” See Olson, E. T. (1997). *The Human animal: Personal identity without psychology* (p. 97). London: Oxford University Press.

[159] While Alvin Goldman first called attention to this problem for reliabilism, Richard Feldman and Earle Conee defend the most robust version of the problem. See Goldman, A. (1979). What is justified belief?; Feldman, R. (1995). Reliability and justification. *The Monist*, *68*, 159–174; and Conee, E., & Feldman, R. (1998). The generality problem for reliabilism. *Philosophical Studies*, *89*, 1–29.

[160] The following argument is similar to one made by Tomas Bogardus, although our application of the generality problem and the emergence reply is unique. See Bogardus, T. Undefeated dualism. *Philosophical Studies*, *165*, 445–466.

[161] A select few have argued that thermostats are conscious. This claim was first defended by McCarthy, J. in his paper. (1979). Ascribing mental qualities to machines. In M. Ringle (Ed.), *Philosophical perspectives in artificial intelligence*. Atlantic Highlands, NJ: Humanities Press. For a response, see Searle, J. (1980). Intrinsic intentionality. *Behavioral and Brain Sciences*, *3*, 450–456.

However, this revision is untenable, as it assumes the very thing under dispute: the truth of physicalism over dualism. Some might cite neurological evidence that the brain and central nervous system interact and correlate with consciousness. But such evidence does not uniquely support physicalism. The neurological data is consistent with dualism. Even ardent physicalists admit that the relevant neurological evidence regarding consciousness does not rule out dualism.[162] Therefore, this emergentist revision of DPA begs the question.

DPA is, therefore, in a seemingly impossible dilemma. Either DPA is so general that it forces us to reject uncontroversial and obviously justified beliefs, or it is restricted by a criterion that the SD can reject. Either result undermines DPA.

5.5.8.4 A Superior Physicalist Account of Self-Awareness

Jaegwon Kim and Paul Churchland have advanced slightly different versions of the most frequently employed physicalist account of self-awareness. Call this *The Scan-Monitoring Approach* (SMA).[163]

> *The Scanning-Monitoring Approach* (SMA): An act of self-awareness is identical to a process of self-monitoring in which one part of the brain (and its associated speak center) discriminatively scans and monitors another part of the brain by way of a causal chain between those parts.

This account assumes externalism about “justification” and a causal account of perception. The only difference between the perception of an external object (an apple) or perception of oneself in self-awareness is that the former involves a proper part of the brain scanning and monitoring an external object. The latter involves a proper part of the brain scanning a spatially internal, different proper part of the brain.

Is an SMA reduction of self-awareness successful? The clear answer is no in light of the arguments earlier in this chapter. Besides the fact that SMA includes no irreducible intentionality nor a *what-it-is-like-for-S* is self-awareness, three other items essential to genuine self-awareness are omitted: *directness*, *identity*, and *subject-simplicity*. Regarding *directness*, genuine self-awareness involves one being *directly aware* of oneself by acquaintance. SMA is mediated by a causal chain that is spatially mediated between the scanning brain part and the scanned brain part. All entities in this process and the whole process itself are entirely bereft of conscious awareness. Regarding *identity*, genuine self-awareness obtains when there is an identity relation between the subject and object of the awareness. But in

---

[162] See, e.g., Corcoran, K., & Sharpe, K. (2006). Neuroscience and the human person. In Crisp, Porter, & Ten Elshof, *Neuroscience and the Soul* (p. 134), and Murphy, N. *Bodies and souls, or spirited bodies?* (p. 112). Cambridge: Cambridge University Press.

[163] See Kim, *Philosophy of mind* (pp. 287–288, 306); and Churchland, *Matter and consciousness* (pp. 120–122).

SMA, the scanning brain part is entirely distinct from the scanned brain part. Both are separable parts. Finally, regarding *subject-simplicity*, in acts of genuine self-awareness, the object of acquaintance is presented as a simple substance. Still, according to SMA, the scanned object will be a complex, spatially extended brain region composed of numerous separable parts. Thus, SMA fails as a reduction of genuine self-awareness.

5.6 A CRUCIAL IMPLICATION OF THE DIRECT SELF-AWARENESS ARGUMENT

Suppose we have a direct awareness of ourselves as simple, immaterial substances. In that case, we have (i) a strong defense of SD, (ii) a plausible and explanatory powerful account of why we have the dualist seemings by unifying them around direct awareness of the self, rather than leaving them as a disparate set of isolated intuitions, and (iii) a good explanation for why dualism is the commonsense view, namely, it expresses what people the world over know to be true based on their direct awareness of themselves.

We have one final point to make. For various reasons, most of what we know about ourselves is epistemically prior to empirical research.[164] And property and substance dualism are not primarily theoretical postulates (and when they are, the range of things they are postulated to explain is distinctively metaphysical and not scientific). However, many hold that while broadly logically possible, SD is no longer plausible in light of advances in modern science, which show that mentality is dependent on the physical. We will give detailed attention to scientific arguments against SD in Chapter 10. For now, we are only drawing attention to a crucial implication of DSA.

Notice that the scientific defeat of SD claim is true only if the justification for SD is primarily its scientific explanatory power given the postulation of the relevant dualist entity.[165] But this depiction is mistaken, or so we have argued. If we are right about this, then physical science is largely irrelevant to debates about SD. The dualist can always appeal to correlations or functional relations between the soul and brain/body. Plausibly, such SD explanations will be empirically equivalent to naturalist explanations, as far as third-person scientific research is concerned with respect to various (two-way) causal and functional dependencies. Moreover, such appeals are not *ad hoc*, given that epistemic justification for SD is prior to science. Indeed, given the arguments of this chapter, it becomes evident that SD must be defeated epistemically prior to employing a claim like this. In this way, the SD argument from direct self-awareness is foundational to the debate about the constitution of human persons and their conscious states.[166]

---

[164] E. J. Lowe, for example, argues that we can have evidence for physical causation only if we have evidence for mental causation. Lowe, E. J. (2008). *Personal agency* (pp. 133–135). Oxford: Oxford University Press. See also, Taliaferro, C. (2018). Substance dualism: A defense. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (p. 48). Oxford: Wiley-Blackwell.

[165] To be clear, DSA does not posit the soul as a scientific explanation of dualist seemings, but a phenomenological and metaphysical model that best explains dualists seemings.

[166] We have chosen to omit treatment of an important argument related to this chapter for dualism, an argument that has disastrous implications for any version of strong physicalism. It is a linguistic argument that focuses on the indexical “I” and related terms (e.g., “my”). The chief advocate of this argument is Geoffrey Madell. See his *The identity of the self* (1981). Edinburgh: The University Press; and *The Essence of the Self.* (2015). New York: Routledge. Cf. Moreland, J. P. (1999). Madell’s rejection of a substantial, immaterial self. *Philosophia Christi*, *1*, 111–114.

PART III

ON THE FUNDAMENTAL UNITY OF CONSCIOUS BEINGS

CHAPTER 6

From Phenomenal Unity to the Synchronic Unity of the Immaterial Self

6.1 INTRODUCTION

The oldest argument for substance dualism begins with the direct first-person awareness of the holistically unified nature of consciousness and concludes that the subject of consciousness is likewise holistically unified, which is not a feature of any physical thing, but a feature of a simple, substantial, immaterial self or soul. For example, in the *New System of Nature*, Leibniz writes,

> Furthermore, by means of the soul or form, there is in us a true unity which corresponds to what we call "I"; this can have no place in artificial machines or in a simple mass of matter, however organized it may be. Such masses can only be thought of as like an army or a flock, or like a pond full of fish, or like a watch composed of springs and wheels. Yet if there were no true substantial unities there would be nothing substantial or real in such a collection.[1]

---

[1] Leibniz, G. W. (1997). New system of the nature of substances and their communication, and the union which exists between the soul and the body. In R. S. Woolhouse & R. Francks (Eds.), *Leibniz's "new system" and associated contemporary texts* (p. 15). Oxford: Oxford University Press. (Originally published in the *Journal des Savants* of 27 June and 4 July 1695), NS1.11.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

Perceptual states—states of consciousness—are, says Leibniz, "the representation of the composite, or what is external, in the simple."[2] This argument has an impressive pedigree beginning with Plato (*Phaedo* and *Theaetetus*) and Plotinus (*Enneads*), defended by numerous early modern philosophers,[3] and has resurfaced in the mid-20th and 21st centuries.[4]

In this chapter, we press several versions of this argument. As we shall see, the unity of phenomenal consciousness presents a significant difficulty for physicalism and panpsychism regarding the possessor of consciousness. Instead, the best explanation for the unity of phenomenal consciousness is a true unity, a substantial soul.

Before moving forward, recall the Husserlian parthood distinction from Chapter 3:

> *Separable Part*: *x* is a separable part of some whole W *iff* (i) *x* is a particular, (ii) *x* is a part of W, and (iii) it is possible for *x* to exist without being a part of W.
>
> *Inseparable Part*: *x* is an inseparable part of some whole W *iff* (i) *x* is a particular, (ii) *x* is a part of W, and (iii) it is not possible for *x* to exist without being a part of W.

Applying the Husserlian distinction to subjects of consciousness, the logical space is exhausted by two theses.

> *Subject Complexity*: Subjects of consciousness are mereologically complex—have at least two separable parts.
>
> *Subject Simplicity*: Subjects of consciousness are mereologically simple—they have no separable parts, are not composed of separable parts.

In naturalist views, the various entities to which a subject (virtual or real) of consciousness is identical—the organism, the brain, an object constituted by an organism, a four-dimensional physical object synchronically composed of its physical/temporal parts—do not have the true unity needed to account for or ground the unity of consciousness because each is a mereological aggregate. Consequently, the naturalist must account for the unity of consciousness within the confines of subject complexity.[5]

---

[2] G. W. Leibniz, *Principles of nature and grace*, in Leibniz, G. W. (1989). *Philosophical essays* (R. Ariew & D. Garber, Ed. and Trans., p. 207). Indianapolis, IN: Hackett. (First published in 1695), section 2.

[3] Other notable examples include Gottfried Leibniz, Thomas Reid, Ralph Cudworth, and Samuel Clark. See Mijuskovic, B. L. (1974). *The achilles of rationalist arguments: The simplicity, unity, and identity of thought and soul from the Cambridge platonists to Kant: A study in the history of an argument*. The Hague: Martinus Nijhoff; Lennon, T., & Stainton, R. (Eds.), *The achilles of rationalist psychology*. Dordrecht, NL: Springer; Rickabaugh, B. (2020). *The conscious mind unified*. Waco, TX: Baylor University; and Textor, M. (2021). *The disappearance of the soul and the turn against metaphysics: Austrian philosophy 1874–1918*. Oxford: Oxford University Press.

[4] See, e.g., Popper, K., & Eccles, J. (1977). *The self and it's brain*. Berlin: Springer-Verlag; Hasker, W. (2009). Persons and the unity of consciousness. In R. C. Koons & G. Bealer (Eds.), *The waning of materialism: New essays* (pp. 175–190). Oxford: Oxford University Press; Swinburne, R. (2013). *Minds, brains, and free will*. Oxford: Oxford University Press; Lund, D. H. (2005). *The conscious self: The immaterial center of subjective states*. Amherst, NY: Humanity Books; Moreland, J. P. (2018). In defense of a Thomistic-like dualism. In J. Loose, A. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 102–122). Oxford, UK: Wiley-Blackwell; Meixner, U. (2004). *The two sides of being: A reassessment of psycho-physical dualism*. Paderborn: Mentis-Verlag; Barnett, D. (2010). You are simple. In R. Koons & G. Bealer (Eds.), *The waning of materialism*. Oxford: Oxford University Press; and Brandon Rickabaugh (2020).

6.2 THE DATUM: PHENOMENALLY UNIFIED CONSCIOUSNESS

6.2.1 Phenomenal Unity

Our unity arguments are restricted to the unity of phenomenal consciousness. A phenomenal conscious state (p-consciousness), ϕ, is such that there is a *what-it-is-like* to be in ϕ. States of p-consciousness include various kinds of experiences, including perceptual, tactile, auditory, and motor sensory as well as cognitive, temporal, and agential. Consider two phenomenal states: (e_1) hearing the ocean and (e_2) seeing the sunset. Each state is qualitatively distinct from the other. Now consider a third phenomenal state, E, hearing the ocean while seeing the sunset. E is qualitatively distinct from both (e_1) and (e_2).[6]

Moreover, E is not reducible to the conjunction of e_1 and e_2 at the same time. Rather, E includes experiencing e_1 and e_2 as one unified state. E has a *what-it-is-like* to hear the ocean while seeing the sunset together as one unified experience. This feature of consciousness is referred to as,

> *Phenomenal Unity*: A state of consciousness E is phenomenally unified when being in E involves jointly experiencing two or more phenomenal states, e_1 and e_2, such that there is a *something-it's-like-to-be-in-both-states-at-once*.[7]

In what Tim Bayne calls an atomistic theory of consciousness, E is composed of "atoms of consciousness"—independent conscious states.[8] The phenomenal states, e_1 and e_2, are separable parts of E. Consequently, E is a mereological aggregate containing and built up by separable parts ("atoms of consciousness"), standing in various external relation-instances which constituted the field's structure—a mereological aggregate of relation instances. At this point one might simply deny phenomenal unity, perhaps due to certain empirical cases. We will address this objection in the final section of this chapter.

---

[5] In Chapter 8, we evaluate Chisholm's idiosyncratic view that the simple self is identical to a rather odd physical entity.

[6] This state is referred to as *objectual phenomenal unity*: when two or more phenomenal states, e_1 and e_2, are experienced in E as being of the same object. Objectual phenomenal unity is challenging to explain neurologically and generates at least two *binding problems*: problems of neurologically explaining how we experience separate pieces of information as bound together and pertaining to the same object.

[7] This is close to what Bayne and Chalmers call *subsumptive phenomenal unity*: all of one's experiences are subsumed within a single, totalizing state of consciousness. We prefer how we have stated holistic unity as the notion of "subsuming" is vague, and as one of us has argued, faces problems. See Moreland, J. P. (2018). Substance dualism and the diachronic/synchronic unity of consciousness. In Loftin, R. K., & Farris, J. R. (Eds.), *Christian physicalism? Philosophical theological criticisms*. Minneapolis, MN: Lexington Books.

[8] Bayne, T. (2010). *The unity of consciousness* (pp. 225–229). Oxford: Oxford University Press.

6.2.2 Phenomenal Holism

Nearly everyone embraces the contrasting thesis according to which the components of the phenomenal field are conscious only as components of that field.[9] More precisely,

> *Phenomenal Holism*: Within a state of phenomenally unified consciousness, E, a single identifiable experience, e_1, is essentially determined by the other experiences, e_2–e_n, synchronically occurring alongside e_1 within E. Consequently, E is more than the sum of its parts, such that facts about the identity and existence of each experience, e_1–e_n, are grounded in facts about E.

In our view, E is a unified whole in which are modes or inseparable parts, e_2–e_n, of either E or much more likely, of the grounding entity or subject of E. As Franz Brentano observes, the phenomenal modalities of E "are neither distinct [separable] things nor parts of distinct [separable] things but belong to a real unity."[10]

6.2.3 Subject Phenomenal Unity

Bayne and David Chalmers also recognize what they call subject phenomenal unity.[11]

> *Subject Phenomenal Unity*: When two or more phenomenal states, e_1 and e_2, are experienced in E as being had by one subject.

The phenomenal states of experiencing the sound of the ocean and the setting sun are experienced as had by a single subject, by me. Subject phenomenal unity is rarely explored. According to Bayne and Chalmers, *subject phenomenal unity* is irrelevant as it is trivially true by definition and tells us nothing about consciousness. This seems false. Many, including Bayne, deny that there is a self or subject that unifies consciousness. And some, like us, think that a simple self or soul is the best explanation for the unity of p-consciousness. Galen Strawson observes the importance of subject phenomenal unity:

> ... the singleness of the total experiential field at t_1 entails the singleness of its subject at t_1, for the material that is experienced is necessarily experienced in a single or unified experiential perspective. The point is a "logical" one, as "trivial" as it is important.[12]

But what kind of subject unity is entailed? Very few think it is subject simplicity, that subjects of phenomenal unity have no separable parts. However, those such as Strawson do not provide an ontology of subject unity.

---

[9] Bayne (2010).

[10] Brentano, F. (1973). *Psychology from an empirical standpoint* (A. C. Rancurello, D. B. Terrell, & L. L. McAlister, Trans., p. 126). London: Routledge. Brackets are ours.

[11] Bayne, T., & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration and dissociation* (pp. 23–58). Oxford: Oxford University Press.

[12] Strawson, G. (2010). Fundamental singleness: How to turn the 2nd paralogism into a valid argument. In P. Basile, J. Kiverstein, & P. Phemister (Eds.), *The metaphysics of consciousness* (Vol. 67, p. 81). Royal Institute of Philosophy Supplement. Cambridge: Cambridge University Press.

6.3 PHENOMENALLY UNIFIED CONSCIOUSNESS CONTRA PHYSICALISM

Briefly stated, physicalist models hold that S is the subject of a phenomenally unified state of consciousness, E, in virtue of either (i) S's separable parts processing E as a whole; (ii) S's separable parts individually processing specific separable parts of E; (iii) the complex structural arrangement of S's separable parts; (iv) the function of S's separable parts; (v) the overlap of S's separable parts. Each view will face specific objections.

6.3.1 The Anti-Distribution Argument

Standard physicalism is committed to the following thesis.

> *Distribution*: It is possible that a phenomenally unified state of consciousness, E, exists distributed across the separable parts of a subject S.

The separable parts might be those of the brain and central nervous system or the whole biological organism. A core premise of most versions of the phenomenal unity argument is,

> *Anti-Distribution*: It is not possible that a phenomenally unified state of consciousness can exist distributed among the separable parts of an object.

The distribution argument is motivated by the following problem. Consciousness seems to be holistically phenomenally unified. In physicalism, consciousness is somehow distributed across the separable parts of a mereologically complex subject. However, distributing a holistically unified state over a myriad of separable parts seems implausible.

Juxtaposing holistic phenomenal unity with the extraordinary mereological complexity of the brain highlights this problem. If phenomenally unified consciousness is distributed, it is so over various subregions of the brain, hundreds of thousands of synaptic inputs of which each synaptic cleft hosts more than 300,000 separable parts, and across one hundred billion neurons and roughly 86 billion cells in total or roughly 37 trillion cells if phenomenal unity is distributed across the organism. The physicalist must explain how phenomenally unified consciousness, which has no separable parts, is somehow distributed across billions of separable parts.

Let's begin with a particular case of distribution. Let the following values hold:

S stands for the subject of a phenomenally unified state of consciousness.
p_1—p_n stands for the separable parts of S.
E stands for the phenomenally unified state of which S is the subject.
e_1—e_n stands for the phenomenal modalities of E.

According to the scattered parts view,

> *Scattered Parts*: S is the subject of E in virtue of S's separable parts individually processing specific sense modalities of E.

The following figure is helpful.

As Figure 6.1 illustrates, the phenomenally unified state, E, is distributed across the neurologically relevant separable parts of S, such that e_1 is processed by p_1, e_2 by p_2, e_3 by p_3, and so on.

As previously explained, p-consciousness is holistically unified (*phenomenal unity holism*) and has no separable parts. However, if E has no separable parts, then there are no such parts of E to distribute. Distribution seems impossible.

One could argue that the inseparable parts of E are distributed across the separable parts of S. In principle, inseparable parts cannot exist apart from their whole. If inseparable parts e_1 and e_2 are in S's separable parts p_1 and p_2, then p_1 could survive the destruction of S, which would also destroy E, which means that e_1 could survive in p_1. Yet, that is not possible, as e_1 could survive the destruction of E, contrary to the inseparability of e_1 from E. So, if E is to be distributed across S, then S must be a genuine whole—one that is not mereologically complex—across which E is distributed. But this is not possible for physicalism as it entails subject complexity. Perhaps one could resist this argument by somehow showing that even though e_1 is found in p_1, it does not follow that e_1 could survive the destruction of S. It isn't clear to us what an account would look like.

Alternatively, one might argue that e_1–e_n are non-phenomenal parts typically correlated with E and are distributed over the separable parts of S or parts of S's brain. Perhaps, on this account, e_1–e_n carry or are themselves packets of information about the content of E. This view avoids the problem of distributing E or E's inseparable parts over S's separable parts. However, this view faces what we call *the process problem*: the problem of explaining how something with no phenomenal properties can process a phenomenally unified state. We will explore this problem below.

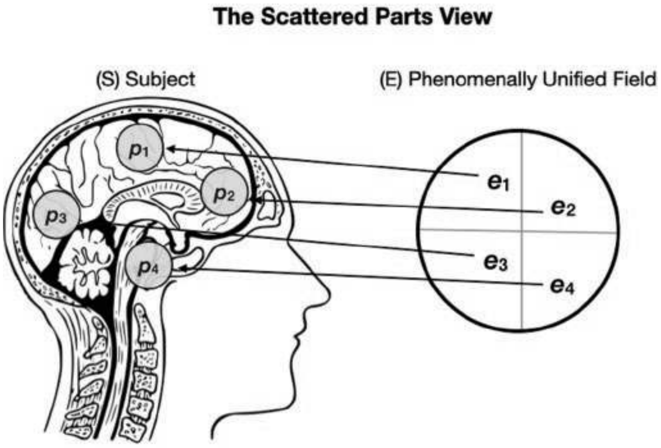


FIGURE 6.1 The Scattered Parts View.[13]

---

[13] Source: Adapted from Pixbay https://ccsearch.creativecommons.org/photos/e78d37a2-53d2-44c2-8c34-644df9400036. Licensed under CC0 1.0.

6.3.2 A Temporal Distribution Problem

We know that the neural processes correlated with consciousness are temporally distributed and, therefore, at least temporally disunified. On average, signals travel along neurons at between 0.5 m/s and 120 m/s, depending on the width and myelination of the neuron. Due to the use of small non-myelin fibers, the neural processing correlated with pain is relatively slow, carrying signals as slow as 2.0 m/s. Due to the use of small fibers without myelin, the neural processing correlated with pain is relatively slow, carrying signals as slow as 2.0 m/s. Additionally, crossing a synaptic cleft between neurons takes at least 0.5 milliseconds. The more neurons involved in visual information processing, the longer the process takes.[14] Therefore, in order for E to be distributed across S, at least some inseparable parts, say e_5–e_{20}, must be in S's separable parts, p_5–p_{20}, which are temporally separated from other parts of E. Again, we have a case of the inseparable parts of E being separated from E in order for E to be distributed across S. This temporal disunity creates the possibility of a temporal gap. Although our perceptions seem unified in one moment (subjectively experienced time), they are in fact temporally scattered (objective time). What we take to be a single experience is the confabulation of a multitude of experiences that are separated by temporal gaps of varying m/s.

This temporal distribution and disunity provide a strong defeater for the possibility of knowledge. C. I. Lewis observes that "All knowledge is knowledge of someone; and ultimately no one can have any ground for his beliefs which does not lie within his own experience."[15] Lewis' student, Roderick Chisholm, connects this insight to the unity of consciousness. The way unified consciousness is presented to us, argues Chisholm, is such that he can be absolutely certain that they are all had by one and the same thing—namely himself. And this is the closest he comes—and can come—to apprehending himself directly. But this awareness that there is something having the properties in question is what constitutes our basis, at any time, for all the other things that we may be said to know at a time.[16] However, if our consciousness is temporally disunified, and presuming that Lewis and Chisholm are correct, then our self-knowledge is not necessarily of one unified self. Temporal disunity makes it possible that what we experience as one individual self is actually a confabulation of several distinct selves that have come and gone, and with it the foundation of knowledge.

One might propose the following to avoid the distribution problem(s).

> *Scattered Wholes*: S is the subject of E in virtue of S's separable parts, each of which processes E as a whole.

In this view, multiple tokens of E are distributed across various non-experiential separable parts of S. Accordingly, S is a subject of phenomenal unity because S's, p_1–p_n, each processes the phenomenal content of E as a whole. This view might avoid the distribution problem(s) insofar as the phenomenal unity of E is somehow preserved.

---

[14] Welsh, T. (2015, June 26). It feels instantaneous, but how long does it really take to think a thought? *The Conversation*. http://theconversation.comit-feels-instantaneous-but-how-long-does-it-really-take-to-think-a-thought-42392. Referenced in Blackmore, S., & Troscianko, E. (2018). *Consciousness: An introduction* (3rd ed., p. 129). New York, NY: Routledge.

[15] Lewis, C. I. (1946). *An analysis of knowledge and valuation* (p. 236). La Salle, IN: Open Court Publishing.

[16] Chisholm, R. M. (1981). *The first person: An essay on reference and intentionality* (p. 89). Minneapolis, MN: University of Minnesota Press.

Although within the logical space, the scattered wholes view is not a plausible form of physicalism unless one gives up certain neurological arguments for physicalism, especially those made from the neurological correlates of consciousness.[17] These studies show that sensory modalities (e.g., those in the visual cortex) are individuated from each other. That is, S has a multitude of individuated brain regions and subregions responsible for distinct sense modalities of E but not of E as a whole.

6.3.3 The Process Argument

Considering the neurological facts just mentioned, *as understood by the physicalist*, the physicalist is plausibly committed to,

> *Process*: It is possible that a phenomenally unified state of consciousness be processed by separable parts that have no phenomenal properties.

Process produces what we will call the *process problem*: the difficulty of explaining how the separable parts of a subject S, which have no phenomenal properties, process the phenomenally unified content of E, such that S is the subject of E. This problem leads to what we call the process argument from phenomenal unity or the *process argument* for short.

Recall that the physicalist is committed to a thesis about the explanation of consciousness such that a complete understanding of the microphysical facts correlated with consciousness will, in principle, suffice for a complete understanding of how a subject has a phenomenally unified consciousness. Plausibly, from this thesis follows a reduction principle.[18]

> *Property Reduction Principle*: If an object O is an ordered aggregate,[19] then all the monadic properties of O are logically or conceptually necessitated by the properties and relations between or among O's separable parts.

Consider another physicalist attempt to explain phenomenal unity: structured parts. On the structured parts view, the parts of E are individually had by various non-phenomenal parts of S, which stand in a complex structure of external relations to each other and S. Structuring the separable parts of S does not help. Facts about the structure of S are reducible to facts about each of S's separable parts.

---

[17] See, e.g., Velmans, M. (2009). *Understanding consciousness* (2nd ed., 15–30). London: Routledge; and Velmans, M. (2007). Where experiences are: Dualist, physicalist, enactive and reflexive accounts of phenomenal consciousness. *Phenomenology and Cognitive Science*, *6*, 547–563.

[18] William Hasker has presented a version of the *process argument* employing this principle inspired by a reductionist principle formulated by Wilfrid Sellars: if *x* is an aggregate (a system of parts), then every property of *x* must consist in facts about *x*'s parts and the relations among them. My principle is more precise and will make possible certain instances of the process argument unavailable to Hasker's version. See Hasker, W. *The emergent self*, pp. 125–135.

[19] As explained in Chapter 3, O is an aggregate *iff* (i) O has at least one separable part, and (ii) the separable parts of O stand only in external relations between each other and O. An ordered aggregate is a functionally unified aggregate, where an object O is functionally unified *iff* (i) O is an aggregate, and (ii) the separable parts of O work together to accomplish some end as a consequence of the external relations among *O*'s parts.

6.3.4 A Non-Phenomenal Processing Problem

In standard physicalism, the separable parts of the brain or organism are not themselves conscious.[20] Consequently, the separable parts of my brain can process only those properties that are neither phenomenal nor phenomenally unified. My brain processes the non-phenomenal sensory information, which becomes conscious only when the brain's separable parts interact in the right way(s). When this occurs, the complex subject experiences phenomenal properties. However, it isn't clear how phenomenal content can be distributed across or processed by p_1–p_n as, via standard physicalism, p_1–p_n have no phenomenal properties. We are owed an account.[21]

6.3.5 A Micro-Subject Problem

Phenomenal content is constituted by the phenomenological details of what it is like for a subject to be in a particular phenomenal state. The content of a phenomenal state is not reducible to non-phenomenal content. Consequently, for p_1–p_n to process the phenomenal content of E, it is not sufficient, perhaps not even necessary, for p_1–p_n to distribute any non-phenomenal content. This poses serious problems.

Recall the *subject necessity* thesis: For all x, if x is a phenomenal state, then necessarily, there is a subject of x. All phenomenal states have a *what-it's-like-to-experience-that-state*. And you only get a *what-it's-like-to-experience-that-state* quality if there is something that does the experiencing or, perhaps better, *has* the experience (the quality), a subject. Combined with distribution, we get the following: If p_1–p_n carry the phenomenal content of E via e_1–e_n, then p_1–p_n are each an individual subject of the corresponding phenomenal state, e_1–e_n. Therefore, if a state of phenomenally unified consciousness is spread out among my separable parts (supposing I have separable parts), then I am not a subject of phenomenally unified consciousness.[22] I am a multitude of subjects, none of which is a subject of E, but only of a sense modality of E.

For the sake of argument, let's suppose that physicalism can explain the distribution and process of a phenomenally unified state. Another problem arises, an argument common among various early modern philosophers. According to the *scattered parts view*, what explains the fact that S is the subject of a phenomenally unified state E has to do with facts about the distribution of e_1–e_n across S's separable parts, p_1–p_n. However, distributing aspects of a phenomenally unified state is insufficient for phenomenal unity.

Given the neurological evidence as interpreted by the physicalist, the consequence is far more dramatic. The separable parts of a subject are exponentially more numerous than sense organs. The problem isn't of an ear hearing and the eyes seeing, but at least one of hundreds of thousands of synaptic inputs across one hundred billion neurons somehow carrying or processing minute aspects of E.

---

[20] If the subregions of my brain are aware of the phenomenal information they process, then other problems emerge.

[21] We can distinguish properties involved in the process of phenomenal unity and properties that are the target of the processing (i.e., the content that is being processed). We've argued that the former is at issue. One might reply that in some physicalist views, functionalism, for example, the latter is the real issue. However, if the content that is being processed is phenomenal content, as it should be, then the parts of E are still scattered over the nonconscious parts of S.

[22] We are assuming that one subject cannot be made up of many, which is common in standard physicalism.

As the fundamental phenomenal unity thesis states, this does not suffice. Phenomenally unified states are not mere collections of sense modalities experienced at *t*. There is an additional phenomenological fact, that of having the experience of all of these sense modalities in a holistically unified way at *t*. There is a what-it-is-like to have the totalizing experience. To avoid the *distribution problem*, one needs a means of holding that I am conscious in a phenomenally unified way via various cognitive mechanisms, none of which are conscious of anything.

6.3.6 The Function Argument

Perhaps the physicalist might suggest that the right functioning of S's separable parts is what makes an aggregate of matter the subject of phenomenal unity. More precisely,

> *Functional Parts*: S is the subject of E in virtue of how S's separable parts function, each of which processes each modality of E.

Like all scattered parts views, the phenomenal parts of E are distributed across the many non-phenomenal parts of S, such that e_1 is exemplified by p_1, e_2 by p_2, e_3 by p_3, and so on. What explains the fact that S is a subject of phenomenal unity has to do with facts about how S's brain functions in virtue of certain parts p_1–p_n realizing various functions.

The function argument is made from the following insight:

> *The Function Problem*: A mereologically complex subject S cannot be holistically unified as the functions of S consist in the functions of S's separable parts, which are only functionally unified.

The function argument depends on the thesis that S's being a subject of E cannot consist of the function of S's separable parts, each of which is not a subject of E.[23] As we understand it, the function an object O has is determined, at least in part, by O's nature. If O is a genuine whole, then O is holistically unified. Recall from Chapter 3:

> *Genuine Whole*: *W* is a genuine whole *iff* (i) *W* is a whole, (ii) *W* has only inseparable parts, and (iii) *W* is unified intrinsically by, and only by, the nature of *W*'s inseparable parts.

Applied to unity we get the following.

> *Holistic Unity*: O is holistically unified *iff* (i) O is a Genuine Whole, and (ii) the parts of O cause some effect as a consequence of the internal relations among O's parts.

However, if O is an aggregate, then O is only functionally unified. Recall,

> *Aggregate*: *x* is an aggregate *iff* (i) *x* has at least one separable part, and (ii) the separable parts of *x* stand only in external relations to each other.

---

[23] Hasker makes a similar argument, although he does not clarify the distinction between something functioning as a whole and something functioning as a system of parts. See, e.g., Hasker, *Emergent self*, p. 124.

> *Functional Unity*: O is functionally unified $_{=\text{def.}}$ (i) O is an aggregate, and (ii) the separable parts of O work together to cause some effect as a consequence of the external relations among O's parts.

If O is an ordered aggregate, then a principle similar to the property reduction principle seems true.

> *Functional Reduction Principle*: For any complex object (one with a plurality of separable parts) O, if O performs function F, then O's F consists of O's separable parts p_1–p_n and subfunctions/activities f_1–f_n, such that p_1 performs f_1, p_2 performs f_2, and so on.

For example, in Leibniz's Mill we are introduced to a mereologically complex machine, the mill, the functions of which are just the pushes and pulls—external relations—of the mill's separable parts. The mill's realizing F consists of the mill's separable parts p_1–p_n realizing sub-functions f_1–f_n, such that p_1 performs f_1, p_2 performs f_2, and so on. This seems obvious, at least concerning additive properties, such as spin, mass, and density.

Now here is the problem. Facts about the functional unity of a subject do not entail facts about the holistic unity of a subject, nor do facts about the functional unity of a subject entail facts about the phenomenally unified consciousness of a subject. The Functional Reduction Principle helps explain why functionalism about mental states is susceptible to Block-style counterexamples.[24]

6.3.6.1 The Function Argument Against Animalism

Consider *animalism*, roughly, the view that we are each a human animal, a biological organism. Notice that without an ontology specifying what an organism is, animalism is compatible with various forms of hylomorphism, idealism, and psycho-physical dualism.[25] Animalism is even compatible with substance dualism or *ensouled animalism*, according to which we are each identical to an animal that has an immaterial soul as a part.[26] One might think part of animalism's attraction is its seeming ability to dodge the hard problem of consciousness and intentionality or the unity of phenomenal consciousness by its ontological ambiguity.

We will, therefore, consider a standard form of animalism defended by Peter van Inwagen and Eric Olson.

> *Biological Animalism*: I am identical to an animal, a complex of atomic simples with a specific complex structure called a life.

Much depends on the ontology of this specific structure, this life. Olson explains,

> By a life I mean ... a self-organizing biological event that maintains the organism's complex internal structure ... an organism's life enables it to persist and retain its characteristic structure despite constant material turnover.[27]

---

[24] Both objections posit counterexamples in the form of creatures or groups of creatures functioning just like human beings, or the brain of a normal human being, while having either no phenomenal states or the wrong phenomenal states. See, e.g., Block, N. (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science*, *9*, 261–325; and Block, N., & Fodor, J. (1972). What psychological states are not. *Philosophical Review*, *81*, 159–181.

[25] E.g., Toner, P. (2011). Hylemorphic animalism. *Philosophical Studies*, *155*(1), 65–81; and Thornton, A. K. (2016). Varieties of animalism. *Philosophy Compass*, *11*(9), 515–526.

[26] Thurow, J. C. (2018). Animals with souls. *Sophia*, *57*(1), 85–101.

[27] Olson, E. T. (2007). *What are we? A study in personal ontology* (p. 28). New York: Oxford University Press.

More specifically,

> *Life Principle*: If the activities of *x*s at t_1 constitute a life, and the activities of *y*s at t_2 constitute a life, then the organism that the *x*s compose at t_1 is the organism that the *y*s compose at t_2 *iff* the life constituted by the activity of the *x*s at t_1 is identical to the life constituted by the activity of the *y*s at t_2.[28]

So, a life is constituted by the activities of an organism's atomic simples at t_n that constitute the organism's life at t_n. The *Life Principle* tells us what unifies a human animal synchronically and diachronically. There is a strong case to be made that animalism cannot account for the endurance of a human animal or organism.[29] In this chapter we will focus on synchronic unity and reserve diachronic unity for the following chapter.

On biological animalism, a group of simples (separable parts) compose a human organism just in case those simples function in the right way. More precisely,

> *Animal Composition*: For any *x* and *y* and *z*, if *x* is an organism and *y* is the life of *x*, then the *z*s wholly and entirely compose *x iff* the activity of the *z*s constitutes the life of *y*, and for any *w*, *w* is one of the *z*s *iff* the activity of *w* partly constitutes the life of *y*.[30]

As it is, the mereologically complex organism is unified insofar as its separable parts count toward sustaining the structure and activities of the life. This is what makes it possible for the terms "organism" and "animal" to express sortal concepts. However, animalists have not made this point clear. In this view, a life consists of the activities of all the parts of an organism at a time. Hence, the unity of an organism is derived from the activities of the organism's parts at a given time.

Now consider van Inwagen's affirmation of the core commitment of the unity of phenomenal consciousness argument: "In my view, we do have a need for 'one,' that is, for the individual thing that thinks. I do not see how we can regard thinking as a mere cooperative activity."[31] On the same page, however, van Inwagen clarifies his view: "things can work together to think only in the sense that they can compose, in the strict and mereological understanding of the word, an object that thinks."[32] But why think that is possible? Why assume that a subject of phenomenal consciousness can be a storm of simples satisfying a particular biological function? A human animal under this construal is only functionally unified, as it is only an aggregate of separable parts (atomic simples). Just as the functions of things working together cannot produce phenomenal unity, neither can the functions of things produce a subject of phenomenal consciousness.

Perhaps the function argument can be avoided by a view in which the separable parts of S's overlap each other.

> *Overlapping Parts*: S is the subject of E by virtue of the overlap of S's separable parts, each of which individually processes a separable part of E.

---

[28] van Inwagen, *Material beings*, p. 145.

[29] See Rickabaugh, B. (work in preparation). Enduring problems for animalism.

[30] Peter van Inwagen holds to a stricter theory of composition, not entailed by his animalism. See his (1990). *Material beings* (p. 91). Ithaca, NY: Cornell University Press.

[31] van Inwagen, P. (1990). *Material beings* (p. 118).

[32] Ibid.

Two things overlap if they have at least one part in common. That is, *x* and *y* overlap ($x \circ y$), just in case there exists an object *z* such that $z \leq x$ and $z \leq y$.

Overlap: $x \cdot y \leftrightarrow \square \exists z [z \leq x \wedge z \leq y]$.

The sense modalities of E are distributed across the various non-phenomenal parts of S, such that e_1 is exemplified by p_1, e_2 by p_2, e_3 by p_3, and so on.

As Figure 6.2 illustrates, the fact that S is a subject of phenomenal unity of E is explained by facts about the overlap of S's separable parts p_1–p_n.

First, plausibly, the overlap of separable parts p_1–p_n forms another separable part of S. And it isn't at all clear how adding another separable part to an aggregated subject is helpful. Rather, it seems to exacerbate the problem.

Second, the overlapping parts view faces problems from neuroscientific evidence. New metanalyses of split-brain studies have invalidated the claim that such cases show a disunity of consciousness.[34] These studies show that unified consciousness may arise and persist without pronounced communication between brain hemispheres. Interestingly, this data challenges leading overlap or holistic theories of consciousness, such as the integrated information theory, global workspace, and the recurrent processing theory.[35] The evidence suggests that there is no significant overlap among at least some neural subregions.

6.3.7 An Emergentist Objection

One might reply that the *functional reduction principle* does not rule out emergent properties. Given the reasonable assumption that supervenience for simple, emergent properties is local (the supervenient simple property obtains and is dependent on what is going on right there at

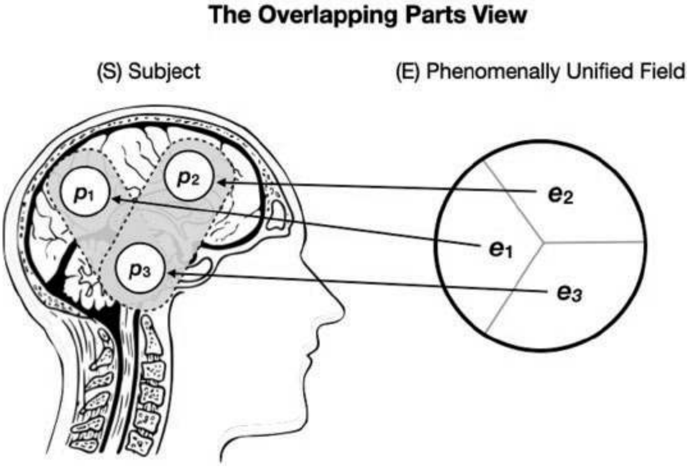


FIGURE 6.2 The Overlapping Parts View.[33]

---

[33] Source: Adapted from Pixbay. https://ccsearch.creativecommons.org/photos/e78d37a2-53d2-44c2-8c34-644df9400036. Licensed under CC0 1.0.

[34] See, e.g., de Haan, E. H. F., et al. (2020). Unified tactile detection and localisation in split-brain patients. *Cortex*, *124*, 217–223; Pinto, Y., et al. (2017). Split brain: Divided perception but undivided consciousness. *Brain: A Journal of Neurology*, *140*(5), 1231–1237; and Pinto, Y., et al. (2017). The split-brain phenomenon revisited: A single conscious agent with split perception. *Trends in Cognitive Sciences*, *21*(11), 835–851.

[35] Pinto et al., The split-brain phenomenon revisited, p. 844.

the subvenient base), the *property reduction principle* disallows emergent properties exemplified by complex objects like O taken as an irreducible whole. But it does not disallow each of the relevant parts of O to have an emergent property as long as these parts are mereological simples.

There are at least two problems with this emergentist objection. First, we are in complete agreement with the localization of emergent properties. After all, it seems to be the specific physical conditions at location l_1 that are responsible for the emergent property to arise. If this is so, then physical factors at other locations are irrelevant except insofar as they alter the physical conditions at l_1, so they are just the right conditions for emergence to obtain. But this would imply that the emergent property would exist and only exist at the location of the conditions that gave rise to and continually sustained the relevant emergent property. If this is correct, then we are back to the problem of there being subjects of phenomenal unity that are a collection of phenomenally conscious subjects. Among other things, this result follows from a combination of the *subject necessity* thesis and the *distribution problem*.

Second, we presented a detailed critique of emergent properties in Chapter 2, and we refer the reader to that material. If that critique is successful, then the appearance of phenomenal consciousness and phenomenal unity signals the presence of a new substance—the subject—and not the presence of emergent properties. And this is just what the substance dualist advocates.

6.3.8 Not All Simple Souls Will Do: Contra Emergent Substance Dualism

Not all SDs agree on what it means for the soul or self to be simple. There are reasons to think that only one of these understandings of the simple soul can explain phenomenal unity. Ned Markosian distinguishes between various ways some entity could be mereologically simple, two of which are important for our discussion:[36]

> *The Pointy View*: x is a mereological simple $=_{\text{def.}}$ x is point-sized (x may have spatial location but not extension). (This may, but need not, be combined with a pointy view of space in which case a simple entity overlaps with and only with an unextended spatial point.)
>
> *The MaxCon View*: x is a mereological simple $=_{\text{def.}}$ x is maximally continuous. An object is maximally continuous if it is spatially continuous and the region it occupies is not part of some larger spatially continuous matter-filled region.

We agree with most substance dualists who adopt The Pointy View, which we explained in Chapter 3 is compatible with our notion of mereological simplicity. However, emergent substance dualist William Hasker and his followers demur and embrace a view of the soul that embraces a MaxCon understanding of simplicity or something very close to it.[37] According to Hasker, the soul is simple with respect to separable parts, but it is spatially extended and, thus, divisible.

---

[36] Markosian, N. (1998). Simples. *Australasian Journal of Philosophy*, *76*, 332–340.

[37] On Hasker's view, the soul, the emergent consciousness is itself a spatial entity: "the volume of space within which the emergent mind exists must be at least sufficient to encompass those parts of the brain with which the mind interacts." Hasker, W. *The emergent self*, p. 192.

We reject Hasker's view of the soul for two main reasons.[38] First, as he admits, his view entails that persons (emergent substantial souls) can be divided or fragmented. This would imply that they could exist as percentages of their original self. In fact, Hasker uses divisibility to explain split-brain phenomena. Hasker explains,

> Put briefly, the solution is that there is indeed an immaterial subject of experience, but one that emerges from a complex physical object, namely the brain and nervous system. Normally what is generated by a brain is a single, well-unified conscious subject, but under special conditions, as seen in commissurotomy and multiple personality, the consciousness can divide or fragment.[39]

But as we have argued in this chapter and others, there isn't sufficient evidence to take persons to be divisible. Moreover, we think there are better solutions to split-brain phenomena.[40]

Second, we take Hasker's view to be problematic for the unity-of-consciousness argument that he champions. That argument is better served by a pointy view of the soul. Here's why. Hasker rightly argues that a composed object (one composed of two or more separable parts) does not have the sort of unity needed to explain or ground the unity of consciousness. But Hasker fails to see that a spatially extended, uncomposed soul faces the unity-of-consciousness problem in much the same way as does a composed object.

To see this, consider a movie screen and take it to be spatially extended, yet uncomposed. In this case, its surface is a primitive unit of two-dimensional extension. That surface may not be composed of a diversity of nonidentical parts, but it does, in fact, contain several nonidentical regions. This becomes evident if we divide the screen arbitrarily into a set of nonoverlapping, square-foot regions that, taken collectively, overlap and only overlap with the entire surface. When a movie is projected onto the screen, there is no region that overlaps with the entire movie. Instead, the entire movie is distributed such that each region of the screen overlaps with only one aspect of what is shown on the entire screen.

Analogously, if a Haskerian soul is postulated to explain the unity of consciousness, it will fail to do so. Each subregion of the extended soul will overlap with and only with one aspect of one's totalizing state of consciousness. Only an unextended simple avoids this problem.

6.4 PHENOMENALLY UNIFIED CONSCIOUSNESS CONTRA RUSSELLIAN PANPSYCHISM

Perhaps the naturalist, following some early modern philosophers, can avoid the unity of consciousness argument by adopting *panpsychism*: the view that fundamental elements of the natural world have conscious experiences.[41]

---

[38] For further objections to emergent substance dualism, see Rickabaugh, B. (2018). Against emergent dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 73–86). Oxford: Wiley-Blackwell.

[39] Hasker, W. Persons and the unity of consciousness, p. 184.

[40] See, e.g., Bayne, T. (2008). The unity of consciousness and the split-brain syndrome. *The Journal of Philosophy, 105*(6), 277–300.

[41] E.g., Pierre Bayle (1647–1706) and Anthony Collins (1676–1729). Cf. Solère, J.-L. (2017). Bayle and panpsychism. *Archiv für Geschichte der Philosophie*, *99*(1), 64–101. Those pressing unity of consciousness argument for SD saw this view as no better off than standard materialism. Cf. Clarke, S. (1738). Four defenses of a letter to Mr. Dodwell. In *The works of Samuel Clarke*. London: James, John, and Paul Knapton.

The most popularly defended version of panpsychism includes a constitutive thesis.

> *Constitutive Panpsychism*: Facts about consciousness at the macro-level are not fundament but are (wholly or partially) grounded in/realized by/constituted of facts about the fundamental micro-conscious level consciousness.

Many couple this thesis with *Constitutive Micropsychism*: all facts are grounded in/realized by/constituted of consciousness-involving facts at the micro-level. This gives us the following account of phenomenal unity.

> *Panpsychist Phenomenal Unity*: The fact that S is a subject of a phenomenally unified consciousness is grounded in facts about combining fundamental micro-conscious subjects.

Rather than solving the phenomenal unity argument, this thesis exacerbates the problem. Or so we will now argue.

6.4.1 The Combination Problem(s)

As a starting point, consider panpsychism's own recalcitrant problem—the *Combination Problem*: the difficulty of explaining how the phenomenal states of fundamental physical entities combine into a holistic, phenomenally unified state of consciousness. Those who press the unity of consciousness argument have always seen this problem. Echoing thought experiments from Plato, Leibniz, and Clarke, consider the often-cited passage from William James.

> Take a hundred of them [fundamental phenomenal states], shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean.[42]

After surveying several kinds of combination problems, ranging from subject problems to unity and awareness problems, Chalmers, a defender of constitutive panpsychism, concludes, "It is fair to say that no proposed solution has yet gained much support."[43] Perhaps the leading defender of panpsychism, David Skrbina, has acknowledged that, since panpsychism employs combinatorial processes every bit as much as strict physicalism, it is no more adequate to explain the holistic unity consciousness possesses.[44]

The combination subject-summing problem is the most troubling combination problem for panpsychism.

> *The Subject-Summing Problem*: It is possible for a whole made of conscious parts not to be conscious itself, not to be a further subject of phenomenally unified consciousness. So how does consciousness in the subject's parts help explain how that subject as a whole is a subject of phenomenally unified consciousness?

---

[42] James, W. (1890). *The principles of psychology* (p. 160). New York, NY: H. Holt and Company.

[43] Chalmers, D. (2017). The combination problem for panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (p. 180). Oxford: Oxford University Press.

[44] Skrbina, D. (2005). *Panpsychism in the West* (pp. 264–265). Cambridge, MA: MIT Press.

Notice, the problem isn't just the separability of subjects but the phenomenality of discrete subjects.

6.4.2 The Privacy Argument Against Panpsychist Phenomenal Unity

The subjective and private nature of consciousness shows that any state of phenomenal consciousness is "shut in its own skin," making it impossible for minds to combine, for subjects sum. According to constitutive Russellian panpsychism, the following is true.

6.4.2.1 Mental-Sharing and Subjectivity

If constitutive panpsychism is true, then,

> *Mental-Sharing*: For at least some phenomenally unified states M, if M is a state of a composite subject S, then M is also a state of one or more of S's component subjects, S_1–S_n.

Mental-sharing conflicts with a much more plausible thesis.

> *Subjectivity*: For all phenomenal states M, if M is a state of S, then necessarily, there is a *what-it-is-like-for-S-to-be-in-M*.

There is an apparent tension between *mental-sharing* and *subjectivity*. Call this the *Privacy Problem*. Experiences are subjective. I experience my mental states as my own. However, if each experience is private, then how can the private experiences of my separable parts (micro-subjects) make me a subject of a phenomenally unified consciousness?

Here is the argument formally stated.

> P1. *Mental-Sharing*: If constitutive panpsychism is true, then for at least some phenomenally unified states M, if M is a state of a composite subject S, then M is also a state of one or more of S's component subjects, S_1–S_n.
>
> P2. *Subjectivity*: For all phenomenal states M, if M is a state of S, then necessarily, there is a *what-it-is-like-for-S-to-be-in-M*.
>
> P3. *Privacy*: For all phenomenal states M, if M_1 is a state of S_1, then necessarily M_1 cannot be a state of any subject other than S_1.
>
> P4. For all phenomenal states M, necessarily, M cannot have more than one subject (from 2 and 3).
>
> P5. Therefore, constitutive panpsychism is impossible (from 1 and 4).

Plausibly, constitutive panpsychism entails *mental-sharing*. Rather than defending (P1), we will reply to panpsychist rejections of *mental-sharing* below. What follows is a defense of *subjectivity* and *privacy*.

Subjectivity is not controversial. Phenomenal consciousness includes not only a *what-its-like* phenomenology but a *what-its-like-for-me* phenomenology. My conscious experience of a red flower, for example, includes what Uriah Kriegel calls the *qualitative character* (e.g., the reddish component; or in our view, a feature of the sensory awareness of the flower: being-appeared-to-redly) and the *subjective character*, the *for-me-ness* of my experience.[45] Just as an experience entails a subject of that experience, phenomenal consciousness entails *subjectivity*. "The reason for calling it a *phenomenal property*," says Richard Fumerton, "is just that fact—that it is phenomenologically present to *me*."[46]

According to *Privacy*, for all phenomenal states M, if M_1 is a state of subject S_1, then necessarily M_1 cannot be a state of any subject other than S_1. By Leibniz's Law, we know that,

> *Mental Property Sharing*: for all phenomenal states M, if two subjects, S_1 and S_2, share M, then, necessarily, any essential property of M, when had by S_1, is an essential property of M when had by S_2.

Consequently, if M has an essential property when had by S_1 that M does not have when had by S_2, then S_1 and S_2 do not share M. We will argue for two such properties: self-presentation and mine-ness necessity.

6.4.2.2 Self-Presentation

Consider the following case.

> *Private Correlations*: Willow is a neuroscientist investigating her own brain. Her device shows the neural happenings in her brain in real time. During one study, she simultaneously experiences a toothache while her device shows unique neural happenings in her brain. Later that day, Alyssia is studying the brain of her colleague with the same equipment and simultaneously experiences the pain of her toothache as she sees neural happenings going on in her colleague's brain that are identical to those that appeared when studying her own brain.

In the first study, Willow infers that a correlation between her brain state and her toothache is probable. Yet, in the second study, it is not at all reasonable for Willow to infer that her pain is correlated with her colleague's brain state.

What explains *private correlations* is that Willow's states of phenomenal consciousness are private and self-presenting. Based on our treatment in Chapter 5, by self-presenting, we have in mind the following.

> *Self-Presenting*: x is self-presenting *iff* (i) x is a constituent of a phenomenal state M, and (ii) x presents itself directly to the subject of M.

Willow stands in the ontological relation of self-presentation between her and her mental states. Notice that even if you could access the contents of my phenomenal state, it would not follow that you share my mental state. Suppose that you've studied my brain to the degree that you can, with a high rate of accuracy, recognize that I am recalling a memory. You would have access to facts about the content of my mental state, that it is a memory. Still, you would not have access to the perspectival content of my phenomenal states.[47] You would not stand in the ontological relation of self-presentation between me and my mental state. Therefore, *privacy* is true.

---

[45] See Kriegel, U. (2009). *Subjective consciousness: A self-representational theory* (pp. 46–52). Oxford: Oxford University Press. To be clear, the focus of qualitative character should be on the nature of the phenomenal state and not on its intentional object. Thus, the relevant qualitative character is not a feature of the flower (redness), but a feature of the sensory awareness of the flower (which is being-appeared-to-redly, and it is not itself colored). In spite of this confusion, we take his point to be relevant to present concerns.

[46] Fumerton, R. (2013). *Knowledge, thought, and the case for dualism* (p. 263). Cambridge: Cambridge University Press.

6.4.2.3 Mine-ness Necessity

Not only are phenomenal states self-presenting and intentionally indexed to a particular subject, but they also present themselves to the subject as necessarily a state of the subject. Take a look at your hand. I suspect that you can conceive of the hand you now perceive as belonging to someone else. This is not the case for your phenomenal state of *being-appeared-to-handly*. One's phenomenal states are, in the words of E. J. Lowe, "necessarily presented to one's consciousness in a distinctive fashion which makes their ownership unmistakable ..."[48] More precisely,

> *Mine-ness Necessity*: For all phenomenal states M, if M is a state of subject S, then M presents itself (is given) to S as necessarily a state of S.

Mine-ness necessity is part of the phenomenological content of *what-it-is-like-for-me-to-be-in-M*. The property of mine-ness is sharable, just as is the property of redness. However, the property instance, stated in the first-person, mine-ness-for-me is not sharable. The indexical nature of a phenomenal state is tied to me as its unique subject. It appears to me as necessarily a state of mine. In Husserl's terms, the primary object of my experience (e.g., a tree itself) is constituted as there for everyone, as "the intersubjectively identical physical thing."[49] An object is intersubjective because it is available to the intentional acts of consciousness by more than one subject. The same is not true of the secondary intentional object of my phenomenal states me as necessarily the subject of my experiencing the tree. My phenomenal states are not presented to me as states that might or could be mine. No, they present themselves as necessarily states of mine.

6.4.2.4 A Neo-Aristotelian Ontology of Phenomenal States

Why think that phenomenal states cannot be transplanted just as legs can? Because phenomenal states are not, in the words of Lowe, "presented to us as *objects* of our awareness, but as *constituents* of it, and for that reason, inalienably ours."[50] That is, part of what makes a phenomenal state what it is has to do with the fact that it is a constituent of me. We depict this ontologically as follows.

Phenomenal states are internally related to their subject. By internally related, we mean, as discussed in Chapter 3,

---

[47] LaRock, E. (2008). Is consciousness really a brain process? *International Philosophical Quarterly*, *48*(2) (190), 208.

[48] Lowe, E. J. (1996). *Subjects of experience* (pp. 190–191). Cambridge, Cambridge University Press.

[49] Husserl, E. (1983). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy—First book: General introduction to a pure phenomenology* (F. Kersten, Trans.). Dordrecht: Kluwer Academic Publishers. (First printed 1913), §151.

[50] Lowe, E. J. *Subjects of experience*, p. 191.

> *Internal Relation*: R is an internal relation between *a* and *b iff* (i) facts about R are grounded in facts about the natures of *a* and *b*, and (ii) necessarily, if R fails to obtain, *a* and *b* are altered.[51]

For example, if the color yellow stands in an internal relation (brighter than) to the color purple, then anything that did not stand in the brighter than relation to purple could not be yellow. So, if a phenomenal state M_1 stands in an internal relation R (self-presenting) to a subject S_1, then part of what makes M_1 the very thing it is that M_1 stands in R to S_1. The identity of M_1 is partly determined by being self-presenting to S_1.

Moreover, a phenomenal state is a modification of the subject, a mode of the subject. The ontology of a mode includes the ontological notion of a property.

> *Property*: *p* is a property *iff* (i) *p* is a universal, and (ii) *p* can be nonspatially in, exemplified by, or possessed by many things at the same time.

As one of us has argued elsewhere, constituent realism is the best view of how properties relate to the ordinary particulars that “have” them.[52] According to *constituent realism*, when an object exemplifies a property, or a collection of objects exemplify a relation, the property or relation (universals) constitutes an immanent essence of the individuated property or relation instances, which in turn, particularize the individual structure in view, keeping it from being an abstract universal. Thus, if the mind exemplifies a mental property, say, the property of being-a-thought-of-London, then that property enters into the very being of the mind, the immaterial self, as a metaphysical constituent.[53]

On standard Aristotelian metaphysics, a mode is a dependent particular, a way or manner in which a thing exists or could exist ontologically prior to its actualization. More specifically,

> *Mode*: *x* is a mode of *y*, *iff* (i) *x* is a property (or set of properties) of *y*, (ii) *x* includes the nexus of exemplification, and (iii) *y* individuates *x*.

By “nexus of exemplification,” we have in mind the realist notion of a tie of predication between a property *x* and an individuator in each property instance of *x*.[54] Accordingly, a phenomenal state is the state of a subject exemplifying a phenomenal property of which it stands in at least one internal relation.

This brief sketch provides an ontology of phenomenal states that is consistent with and given to us in phenomenological reflection. Moreover, it explains the *mine-ness necessity* of phenomenal states. Phenomenal states are private because part of what makes a phenomenal state what it is has to do with the fact that it is a constituent of me. No one else can stand in this relationship to my phenomenal states.

In summary, phenomenal states are self-presenting. In this mode of self-presentation, my phenomenal states are intentionally related to me and only me. In this mode of self-presentation, my phenomenal states present themselves as necessarily states of mine. These properties are dependent on the specific subject that has them. Moreover, these properties are partly constitutive of any phenomenal state. Therefore, *mental-sharing* is false; phenomenal states cannot be phenomenally unified into a subject of a phenomenally unified conscious state. The constitutive panpsychist solution to the phenomenal unity problem fails.

---

[51] This is distinct from *external relation*: *R* is an external relation between *a* and *b iff* (i) facts about *R* are not grounded in facts about the natures of *a* and *b*; and (ii) possibly *R* fails to hold between *a* and *b*, while at least *a* or *b* remains unaltered.

[52] Moreland, J. P. (2013). Exemplification and constituent realism: A clarification and modest defense. *Axiomathes*, *23*(2), 247–259.

[53] See, e.g., Willard, D. (1999). How concepts relate the mind to its objects: The “God’s eye view.” *Philosophia Christi*, *1*(2), 5–20.

[54] See Moreland, *Universals*, pp. 103–112.

6.4.3 Objections to the Privacy Argument

What premises of the privacy argument can be avoided? Denying (P2)—*subjectivity*—is not going to help. The point of constitutive panpsychism is that the combination of conscious fundamental parts avoids the physicalist's problem of explaining subjectivity. What of the other premises and other independent objections?

6.4.3.1 Deny Mental-Sharing

Perhaps one could deny the first premise that at least some experiences belonging to a composite subject must also belong to one or more of their components. Pierfrancesco Basile, one of the only panpsychists who deny *mental-sharing*, argues that although we lack a firm understanding of phenomenal parts and wholes, we do not know that mental composition is impossible.[55] Panpsychism supposedly avoids combination problems by denying that smaller minds constitute minds like ours in the same way that chairs or mosaics are constituted by wood or stones.[56] This isn't to reject constitutive Russellian panpsychism but to make room for an alternative understanding of constitution.

Basile's reply is not a rejection of *mental-sharing*, but only the claim that we might have a reason to reject *mental-sharing* provided we develop the mereology required for doing so. Basile's objection is an optimistic version of the mereological mystery objection. This optimistic attitude is justified only if there are strong reasons to think we will discover such a mereology. Basile offers none.

Goff rejects *mental-sharing* by arguing that a macro-subject S does not, as a whole, have the experiences of its conscious parts. In Goff's view, although the experiences of S's conscious parts *constitute* S's experiences, they *do not characterize* S's experiences.[57] This view assumes that the phenomenality of S's micro-phenomenal parts is experientially irrelevant to the phenomenality of S's experience. If this is the case, then what motivation is there for panpsychism in the first place? We needn't posit conscious parts, but only parts with powers that, when combined in the right way, constitute a subject of consciousness. Goff's rejection of *mental-sharing* provides a reason for rejecting constitutive Russellian panpsychism. Rejecting *mental-sharing* is, therefore, not a promising move.

6.4.3.2 Reject Privacy

The only option left is to reject privacy, which Roelofs prefers.[58] The challenge is to do so in a way that plausibly retains the subjective character of phenomenal consciousness. Can privacy be separated from phenomenal consciousness? Can privacy be separated in such a way as to retain what the constitutive panpsychist needs as an account of phenomenally unified consciousness? One way to do this is to deflate the subjectivity of the micro subjects that constitute the macro subject S. This strategy assumes that deflating the micro subject is enough to make it possible for their phenomenal states to be shared. This seems obviously false. The macro subject S cannot be deflated. S's phenomenal states are, therefore, private and cannot be the state of anything other than S. The problem is on the receiving side of the sharing relation. Deflating the sharer does not diminish the privacy of the one who should receive the phenomenal state.

---

[55] Basile, P. (2010). It must be true—But how can it be? Some remarks on panpsychism and mental composition. *Royal Institute of Philosophy Supplement*, *67*, 93–112.

[56] See also, Basile, P. (2009). Back to Whitehead? Galen Strawson and the rediscovery of panpsychism. In D. Skrbina (Ed.), *Mind that abides* (pp. 179–199). Amsterdam: Benjamins Publishing.

[57] Goff, P. (2017). *Consciousness and fundamental reality* (p. 194ff). Oxford: Oxford University Press.

[58] Most panpsychists are unwilling to give up privacy. See, e.g., Basil. It must be true, pp. 111–112.

Perhaps one could reject privacy by appealing to empirical cases of *mental-sharing*. The most plausible case would be that of Krista and Tatiana Hogan, craniopagus conjoined twins who share a brain structure consisting of a neural bridge between their thalamus, which regulates consciousness, sensory and motor signals.[59] In virtue of this neural bridge, one brain receives signals from the other brain and vice versa. Although the twins have not undergone rigorous neurological study, it seems as if one can receive some of the neuronal sensory inputs of their twin's brain that are correlated with taste and sight, in addition to their own neuronal sensory inputs. Is this a case of *mental-sharing* or an absence of *privacy*?

It is clear from their phenomenological reports that Krista and Tatiana can tell their own sensations apart from their twin's. Krista's sensations are experienced as hers from her unique first-person perspective, and the same is true for Tatiana. Krista doesn't report having Tatiana's experiences. Instead, Krista reports having access to Tatiana's neuronal sensory inputs. Neither twin has a shared access consciousness or phenomenal consciousness, although it appears that they can receive some of each other's neuronal sensory inputs, which they distinguish phenomenologically. The phenomenological distinctness of each twin's mental state entails that the twins do not share phenomenal states.

Not only is there evidence that we experience our mental states as necessarily our own, but our mental states also present themselves in such a way that we can distinguish between present states and past states being experienced at the present. Wilder Penfield, a pioneer in modern neurosurgery, noted that patients who were conscious while undergoing neurosurgery were aware that the sensation from their memories was being induced by brain stimulation to them but not by them. In one case, a patient of Penfield's was startled when he realized he was having an experience of laughing with his cousins on a farm in South Africa and, at the same time, was also conscious of being in the operating room in Montreal.[60] He could clearly distinguish experience in surgery from the experience evoked by the surgeon.

Lastly, some dualists make arguments (e.g., introspection and knowledge arguments) against physicalism from the fact that mental states are subjective and introspectively available, while physical states are not.[61] One might see the argument of this chapter falling into such a category. A common objection, one repeatedly made by Paul Churchland, is that these arguments error by moving from an epistemological premise about subjectivity to a metaphysical conclusion about the distinctness of mental from physical states.[62] Roelofs makes the same mistake in his rejection of privacy.[63] However, as we argued in Chapter 4, these arguments fail in multiple ways. Accordingly, we think John Searle got it right when he observed that

---

[59] Dominus, S. (2011, May 25). Could conjoined twins share a mind? *New York Times*. Permalink: http://nyti.ms/19p4PMI. We are grateful to Alexander Pruss for raising this concern.

[60] Penfield, W. (1975). *Mystery of the mind: A critical study of consciousness and the human brain* (pp. 55–56). Princeton, NJ: Princeton University Press.

[61] Some argue that introspection supports property dualism, while others argue that it supports substance dualism. For recent examples, see: Fumerton, R. (2013). *Knowledge, thought, and the case for dualism*. Cambridge: Cambridge University Press; and Robinson, H. (2016). *From the knowledge argument to mental substance: Resurrecting the mind*. Cambridge: Cambridge University Press.

> The point of [these arguments] is that there exist real phenomena that are necessarily left out of the scope of their [e.g., Mary's] knowledge, as long as their knowledge is only of objective, third-person, physical facts ... The point about the epistemology, the information, is just a way of getting at the underlying ontological difference.[64]

Positing that subjects of phenomenal unity are explained by these subjects having conscious separable parts in no way helps the naturalist avoid the subject phenomenal unity problem. The privacy argument against constitutive Russellian panpsychism shows that phenomenal states of one subject cannot constitute the phenomenal states of another subject. Phenomenal states are uniquely tied to their subject. For the subject of phenomenal unity, a great host of phenomenal states are tied to the subject in a phenomenally unified way.

6.5 GENERAL OBJECTIONS TO THE UNITY OF CONSCIOUSNESS ARGUMENT

6.5.1 Rejecting Phenomenal Unity

One might appeal to empirical studies (e.g., split-brain experiments) to deny that p-consciousness is unified. In brain bisection surgeries, the corpus callosum, the myelinated sheath responsible for communication between the brain's right and left hemispheres, is severed, resulting in a "split brain." Under certain contrived experimental circumstances, split-brain patients seem to have two centers of consciousness that can be cut off and even at odds with each other. From this some infer that p-consciousness is not necessarily unified or unified at all.

This objection faces several problems. First, at least three recent meta-analyses and empirical studies undermine the divided centers of consciousness interpretation of split-brain cases.[65] These studies show that unified consciousness may arise and persist without pronounced communication between brain hemispheres.

---

[62] See, e.g., Churchland, P. (2013). *Matter and consciousness* (3rd ed., pp. 50–52). Cambridge, MA: MIT Press; Churchland, P. (2013). Consciousness and the introspection of "qualitative simples." In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience* (pp. 35–56). New York, NY: Springer; and Churchland, P. (1985). Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy*, *82*, 8–28.

[63] Roelofs, L. (2016). The unity of consciousness, within and between subjects. *Philosophical Studies*, *173*, 3199–3221.

[64] Searle, J. (2004). *Mind: A brief introduction* (p. 97). Oxford: Oxford University Press. Cf. 96–99.

[65] Pinto, Y., et al. (2017). Split brain: Divided perception but undivided consciousness. *Brain: A Journal of Neurology*, *140*(5), 1231–1237; Pinto, Y. et al. (2017). The split-brain phenomenon revisited: A single conscious agent with split perception. *Trends in Cognitive Sciences*, *21*(11), 835–851; and de Haan, E. H. F., et al. (2020). Unified tactile detection and localization in split-brain patients. *Cortex*, *124*, 217–223.

Second, in these cases, phenomenal consciousness remains experienced as holistically unified. Third, nearly everyone working on the unity of consciousness finds it difficult, if not impossible, to eliminate or reduce phenomenal unity, even though doing so would best serve their commitment to physicalism.[66] Last, this objection arises by adopting the total phenomenal unity thesis: necessarily, all of *S*'s phenomenal states at *t* are phenomenally unified.[67] While the total phenomenal unity thesis might be true, our unity argument does not require it. All we argue is that if a subject has a phenomenally unified consciousness, then that subject is mereologically simple.

6.5.2 Tim Bayne's Explanatory Failure Objection

Until recently, the most detailed objection to the unity of phenomenal consciousness argument was made by Warner Shrader.[68] One of us has provided detailed replies to Shrader.[69] We will, therefore, focus on an objection made by Bayne, what we call *the Explanatory Failure Objection*. Against Leibniz's unity of consciousness argument, Bayne argues that inferring subject simplicity from the failure of physicalism is unmotivated as it offers no explanation of consciousness.[70] Bayne assumes that unity of consciousness arguments succeeds or fail based on their ability to explain the unity of consciousness better than naturalism. But why suppose that is true?

The defender of subject complexity should recognize that subject simplicity does not face the subject phenomenal unity problems and attempt to reject subject complexity on other grounds. As Geoffrey Madell urges, defenders of the phenomenal unity argument need only recognize that in their view, "there is no special problem about what it is for a group of experiences occurring simultaneously to belong to the same mind."[71] It is a mistake to think the subject phenomenal unity argument posits subject simplicity as an explanation of subject phenomenal unity. The arguments we've raised against subject complexity, in turn, show that simplicity is true. The argument is a simple disjunction. If the logical space is exhausted by subject simplicity and subject complexity, then arguments showing the impossibility of subject complexity are arguments for subject simplicity.

Although not required for the argument to succeed, we will briefly present our substance dualist ontology/explanation of the unity of consciousness. Our SD explanation makes use of several ontological notions already introduced. Here are two more.

> *Substance*: *x* is a substance *iff* (i) *x* is an essentially characterized particular; (ii) *x* has (and is the principle of unity for its) properties but is not had by or predicable of something more basic than it; (iii) *x* is an enduring continuant; (iv) *x* has inseparable parts but is not composed of separable parts; (v) *x* is complete in species.

---

[66] Schechter, E. (2018). The unity of consciousness. In R. J. Gennaro (Ed.), *The Routledge handbook of consciousness* (p. 369). New York, NY: Routledge.

[67] Bayne and Chalmers, What is the unity of consciousness?, p. 33.

[68] Shrader, W. (2006). The unity of consciousness: Trouble for the materialist or the emergent dualist? *Faith and Philosophy*, *23*(1), 33–44.

[69] See Moreland, Substance dualism and the unity of consciousness, pp. 190–193.

[70] Bayne, Problems with unity of consciousness arguments for substance dualism, pp. 208–225. For more on this objection, including Bayne's statement of it, see Chapter 10.

[71] Madell, G. (2015). *The essence of the self: In defense of the simple view of personal identity* (p. 45). New York, NY: Routledge.

> *Immaterial Self*: An immaterial self or soul is (i) a substance; (ii) has no separable parts; (iii) metaphysically indivisible in being (though it may be fractured in functioning); (iv) not spatially extended (though some characterizations hold that it may be spatially located); (v) essentially characterized by the actual and potential properties of consciousness.

The nature of the unity of phenomenal consciousness and the nature of its ground is easily explained. The immaterial self (soul, I) is spatially unextended and not composed of separable parts. The self's having a particular mental state, say, an awareness of a table, occurs when the self exemplifies the property *being-an-awareness-of-the-table*, and thereby forms a mode of the self, namely, *the-having-of-P-by-the-self*. This mode, the self's *being-appeared-to-tablely*, is an inseparable part of the self, which stands in an internal relation to the self. This is what a particular phenomenal state is. This also explains *mine-ness necessity*. That which has conscious states is the simple self, and the "having" amounts to the self's exemplifying various properties of consciousness to form modes of the self.

Synchronically, the various mental modes of the self are unified into one holistically unified phenomenal mode (state) by being modes of the same simple immaterial self and by being internally related to that self. The immaterial simple self has the power of self-awareness and self-reference. This explains subjectivity and the first-person perspective.

Moreover, the immaterial self exemplifies and unifies various properties of consciousness that, in turn, have the property of intentionality. In this way, the self/I has an irreducible, unified totalized conscious state/mode that is about intentional objects. Because modes are inseparable parts internally related to their subject, this totalizing state is unique to one subject and cannot be shared.

Finally, the immaterial self is a unique kind of substance in that it has the power of self-awareness and self-reference. On this view, s*ubject phenomenal unity* occurs when all of one's phenomenal states are modes of the same, simple I. *Objectual phenomenal unity* is explained by saying that two or more p-conscious states are experiences as being of the same object (e.g., the color and shape of a table) because they belong to the same object. Lastly, holistic phenomenal unity obtains because it is the same self that exemplifies the conscious property constituting each phenomenal mode/state, and the self's unification of these into a totalizing mode is due to the self's mereological simplicity.

CHAPTER 7

Mereological Essentialism and the Diachronic Endurance of the Soul

Aside from skeptical worries, which are impossible to live with, it's difficult to imagine an argument that could convince one that they are not identical to the person who began reading this book or even this sentence. That we endure over time seems uncontestable. Many have thought the fact of our own endurance is best, if not only, explained by substance dualism. This much we will argue in this chapter.

While longtime SD defender Charles Taliaferro takes previous versions of this chapter's central argument for SD to be sound, nevertheless, he believes the argument rests on a controversial account of diachronic identity—*mereological essentialism* (ME)—that takes us into murky philosophical waters that are best to avoid.[1] Sociologically speaking, Taliaferro may be right. But we believe a successful argument of this sort can be made, and in this chapter, our aim is to present and defend such an argument.

To secure this aim, this chapter proceeds as follows. Section 7.1 provides a statement of the argument, a clarification of certain premises, and a presentation of different notions of simplicity. Section 7.2 proffers a detailed analysis of the standard mereological hierarchy (a.k.a. layered model of reality, or at least, of the natural world) which constitutes a set of broader worldview considerations in terms of which defenses and criticisms of the mereological argument must be advanced. Section 7.3 contains a presentation of important arguments for the claim that we as human persons sustain strict philosophical identity through a loss or gain

---

[1] Taliaferro, C. (2018). Substance dualism: A defense. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (p. 53). Oxford: Blackwell.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

of separable body or brain parts and offers an evaluation of various defeaters for this claim. The section provides epistemic grounds for believing that we are enduring continuants, viz., simple immaterial substances. In Section 7.4, we seek to establish the fact that if, according to naturalism, we were ordinary-sized macro-physical objects—bodies, brains, subregions of brains (understood on standard naturalism[2])—then we would be mereological aggregates (MAs) and, consequently, would not be enduring continuants. Along the way, we provide arguments against different persistence accounts of diachronic personal identity, sometimes called complex views. In the final section, 7.5, we evaluate two views of organisms that claim to provide an account of human beings as enduring continuants without appealing to a simple, substantial soul: the atomic simple view and the dynamic systems depiction of organisms. Animalism was evaluated in Chapter 6, so it will not be included in this chapter.

7.1 STATING AND CLARIFYING THE ARGUMENT

What should we make of the diachronic identity or persistence of the human person?[3] Are we continuants that remain literally the same through accidental change, especially through change in body parts? We have very deep intuitions that we are literal continuants. As we argued in Chapter 5, this is a properly basic belief grounded in direct self-awareness. For example, the simple act of attending to oneself humming through a tune is such that the literal continuity of the self is made evident, and it is the self that unifies each aspect of humming the tune into the experience of one, single subject.[4] In Section 7.3, we will examine this and other arguments for this commonsense view of Folk Ontology.

Now nearly everyone these days wants to avoid ME, at least for living organisms. Most versions of substance dualism provide a fairly straightforward way of grounding human persons (and animals) as substantial continuants while avoiding mereological problems regarding an organism's body, in this case, the human body: We are simple, immaterial wholes and not MAs. Our "persistence" conditions are different from those of our bodies (or our psychology), and the fact—if it is a fact—that ME applies to our bodies does not affect us. SD also solves certain puzzles—e.g., Tibbles/Tibs—of material constitution, but that is not our concern here.

7.1.1 The Mereological Argument

The argument for a simple substantial enduring soul may be put this way:

1. If something is a physical object composed of separable parts (an MA), it does not endure over time as the same object if it comes to have different parts.[5]
2. My (a human person's) body and brain are physical objects composed of separable parts.

---

[2] Setting aside some versions of staunch hylomorphism, idealism, and psychophysical dualism.

[3] It is customary to use "persistence conditions" in the literature on diachronic personal identity. But "persistence conditions" is ambiguous and can refer either to diachronic Leibnizian identity or, more often, to a weaker relation, e.g., similar to some degree, perhaps sortal dependent, of object-stages standing in some relation R. See Gasser, G., & Stefan, M. (Eds.). (2012). *Personal identity: Complex or simple?* Cambridge: Cambridge University Press.

[4] For a phenomenological defense of this claim in Husserl's writings, see chapter nine, 7–9, 11–12, 58–63. Primary sources are provided in the notes for those pages.

[5] (1) can also be stated modally: If *x* is a physical object composed of separable parts (an MA), it is not the case that, possibly, *x* exists and is composed of different separable parts.

3. Therefore, my body and brain do not endure over time as the same object if they come to have different separable parts.
4. I (a human person) do endure over time as the same object.
5. Therefore, I am not my body or brain.
6. I am either a simple enduring soul or a body or a brain.
7. Therefore, I am a simple enduring soul.

7.1.2 Simplicity

Ned Markosian argues that to understand simplicity, we must start with The Simple Question:[6] What are the necessary and sufficient conditions for some *x* to be simple? Markosian notes that in asking this question, we are not looking for a conceptual analysis of mereological simplicity (an object without proper—separable—parts); rather, we want a nonanalytic principle for some *x* to fall under this concept. Here are some answers to The Simple Question relevant to our concerns. Entity *x* is a mereological simple *iff*:

(i) *The Pointy View*: *x* is point-sized (*x* may have spatial location but not extension).[7]

(ii) *The Metaphysically Indivisible View*: *x* is such that it is not metaphysically possible to divide.

(iii) *The MaxCon View*: *x* is maximally continuous. An object is maximally continuous just in case it is spatially continuous and the region it occupies is not part of some larger spatially continuous matter-filled region.

To these we add a fourth view:

(iv) *The Bare Particular View*: For all entities *e* (separable/inseparable parts, properties/property-instances, relations/relation-instances, events/processes), *e* is not an intrinsic constituent of *x*. There is no metaphysical differentiation within or intrinsic to *x*.

It might seem that (i) entails (ii), but this is not so. Among the late Medieval Aristotelians, Duns Scotus and others who followed him held to an Extensionless Part View of prime matter according to which it is extensionless and, totally apart from added quantity, has substantial parts.[8] However, (ii) would seem to entail (i). If something is metaphysically indivisible, then it is necessarily the case that it is spatially unextended. To get (ii) we must add another necessary condition such that they jointly entail (ii): *x* cannot be composed of separable (e.g., substantial) parts. (iii) raises questions about stuff, atomless gunk, and masses. We will discuss these later in this chapter. (iii) captures William Hasker's view of the soul.[9] Finally, while we believe in bare particulars as individuators of substantial forms taken as universals, they are not directly relevant to the argument above.

---

[6] Markosian, N. (1998). Simples. *Australasian Journal of Philosophy*, *76*, 332–340. Cf. Hudson, H. (2007). Simples and gunk. *Philosophy Compass*, *2/2*, 291–302. We set aside different answers to The Simple Question, e.g., the brutal-composition view, that we deem less relevant.

[7] This may, but need not, be combined with a pointy view of space in which case a simple entity overlaps with and only with an unextended spatial point.

[8] Pasnau, R. (2011). *Metaphysical themes: 1274–1671* (pp. 57–60). Oxford: Clarendon Press.

[9] Hasker, W. (1999). *The emergent self*. Ithaca, NY: Cornell University Press.

In the argument, when we use simplicity or simple, we mean these in sense (ii)—extensionless and not composed of separable parts.[10]

7.1.3 Clarification of the Argument

Premise (1) is the claim that MAs are subject to mereological essentialist constraints and are inconstant objects through a change in separable parts (*sans* eliminativism). (2) expresses the canonical scientific way of construing bodies and brains, especially given the standard mereological hierarchy and the part-priority (usually, micro-physical part-priority) ubiquitously associated with it.[11] Moreover, given that the body and brain are MAs and, therefore, objects, debates about the Doctrine of Arbitrary Unattached Parts are not relevant to them and our employment of (2).

Premises (3) and (4) of the Mereological Argument utilize the classical notion of endurance:

> *Endurance*: An object *x endures* over time *iff* (i) *x* retains strict, absolute, Leibnizian identity through accidental intrinsic change, (ii) *x* is wholly present at each moment at which it exists, and *x* has no fundamental temporal parts.[12]

"Persistence" can be used as a generic claim about an object existing over time, with endurance and perdurance as different ways to spell out persistence. But often, persistence is a weaker claim than endurance in that if an object persists over time, it does not sustain absolute identity.

There are two main versions of persistence in this latter sense with different varieties of each:

> *Perdurance*: An object *x perdures* over temporal interval T just in case (i) there is a series of temporal parts such that all of the series constitute T, (ii) *x* has distinct part-stages (time slices) that exist wholly at each temporal moment constituting T, (iii) the part-stages all stand to each other is some relevant relation R, and (iv) *x* is identical to the relevant history, namely, to the total of all the part-stages standing in R to each other. Objects are space-time worms and exist at different times by having distinct parts at those times.
>
> *Series-of-Instantaneous-Entities view*: An object *x* persists throughout a time-interval T constituted by temporal instances t_i–t_n *iff* (i) at each temporal instant t that constitutes T, there exists one and only one non-repeatable *x*-like entity, (ii) the *x*-like entities form a unified series of instantaneous *x*-like entities S by standing in relation R to each other, and *x* is constituted by S.

Here, an object persists in virtue of there being the relevant series of whole entities at each relevant temporal instant. According to some scholars, this is the view of Theravada

---

[10] See Chapter 3 for a detailed discussion of parthood.

[11] We hold to a view within an Aristotelian family of alternatives (though our argument in this chapter is on behalf of general SD) and, so, do not agree with this depiction of the brain and body.

[12] Some would reject (ii) and allow for either nonfundamental temporal instances or two fundamental entities—temporal instances and enduring wholes. See Koons, R. C., & Pickavance, T. (2017). *The atlas of reality* (p. 539). Malden, MA: Wiley-Blackwell.

Buddhism according to which the "self" is identical to a series of instantaneous I's. More plausibly, this view is eliminativist with respect to the self.

Our justification for premise (1) in the Mereological Argument will be given in Sections 7.2 and 7.4. Premise (2) is an obvious scientific fact, given the depiction of the body and brain as MAs and the standard mereological hierarchy (with micro-physical part priority) as a background constraint, along with the combinatorial processes of the Grand Story and the Location Problem covered in Chapter 2. We will examine different views of organisms and their organs in Section 7.5. Premise (3) states that one's body and brain do not endure but, rather, persist. Premise (4) will be defended in Section 7.3. Strictly speaking, (6) does not present an exhaustive dilemma, but it does exhaust the live options for those staunch naturalists and strict physicalists—whether naturalists or theists—against whom we are arguing. Further, the term "simple" in (6) and (7) means "is spatially unextended and not composed of separable parts."

7.2 THE STANDARD MEREOLOGICAL HIERARCHY OR LAYERED VIEW OF THE NATURAL WORLD

As we mentioned in Chapter 2, the Standard Mereological Hierarchy (hereafter, the MH) is the view naturalists should adopt, and staunch naturalists and strict physicalists recognize that.[13] The MH fits best with the naturalist epistemological commitments and Grand Story (constituted by combinatorial processes and explanations) of how all things came about. Specifically, the individuals, properties, and relations in the MH are easily located in the Grand Story and they best conform to epistemic simplicity—an epistemic value that plays a significant role in defending staunch naturalism and strict physicalism—in comparison to systems or substances in the category of individual, emergent properties in the category of property and internal relations in the category of relation.

As Jaegwon Kim notes, any deviation from the MH and the strict physicalism it entails comes at a cost. [Strict] physicalism cannot be had on the cheap."[14] Thus, Terence Horgan asserted that "in any metaphysical framework that deserves labels like 'materialism', 'naturalism', or 'physicalism', supervenient facts must be explainable rather than being *sui generis*.[15]

And D. M. Armstrong goes so far as to admit that

> I suppose that if the principles involved [in analyzing the single all-embracing spatio-temporal system which is reality] were completely different from the current principles of physics, in particular if they involved appeal to mental entities, such as purposes, we might then count the analysis as a falsification of Naturalism.[16]

---

[13] For good summaries of the standard mereological hierarchy, see Kim, J. (1998). *Mind in a physical world* (pp. 15–19). Cambridge, MA: MIT Press; Inman, R. (2018). *Substance and the fundamentality of the familiar* (pp. 75–113). New York, NY: Routledge.

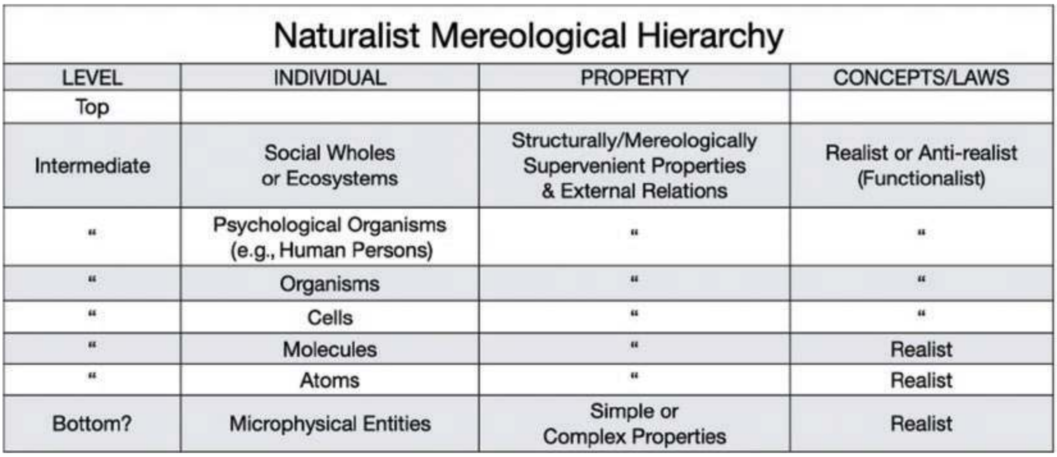
[14] Kim, p. 120.

[15] Horgan, T. (1993). Nonreductive materialism and the explanatory autonomy of psychology. In S. J. Wagner & R. Warner (Eds.), *Naturalism* (pp. 313–314). Notre Dame, IN: University of Notre Dame Press.

[16] Armstrong, D. M. (1978). Naturalism: Materialism and first philosophy. *Philosophia, 8*, 262.

7.2.1 A General Depiction of the Mereological Hierarchy

If the reader disagrees with these assertions, it would be helpful to learn what is wrong with the arguments for it in Chapter 2. In any case, many naturalists, perhaps the majority, adopt the MH, so it will be relevant for them to place our argument in this chapter against the MH as a backdrop. Figure 7.1 gives a general depiction of the MH.



Naturalist Mereological Hierarchy

LEVEL	INDIVIDUAL	PROPERTY	CONCEPTS/LAWS
Top			
Intermediate	Social Wholes or Ecosystems	Structurally/Mereologically Supervenient Properties & External Relations	Realist or Anti-realist (Functionalist)
“	Psychological Organisms (e.g., Human Persons)	“	“
“	Organisms	“	“
“	Cells	“	“
“	Molecules	“	Realist
“	Atoms	“	Realist
Bottom?	Microphysical Entities	Simple or Complex Properties	Realist

FIGURE 7.1 Naturalist Mereological Hierarchy.

There are several pertinent observations to make about the MH. Let level *n* be the bottom level (assuming there is one) and all levels *n*+1, *n*+2, ..., be higher levels.

7.2.2 The Category of Individual

Assuming the level *n* is fundamental and basic, individuals at that level are uncomposed units, usually taken to be atomic simples (having no proper separable parts). All entities at higher levels are MAs. In a later section, we will consider the controversial issue about organisms, along with systems, and what to do with them.[17] An MA is exhaustively decomposable without remainder (except, perhaps, for the loss of an aggregate's surface boundary) into entities at the next lowest level: the parts of the aggregate, along with their properties and the relations between and among them. Given that this decomposition relation seems to be transitive, we end up with some version of microphysicalism to be clarified below. Base-level entities are fundamental; higher level entities are derived.

7.2.3 *The Category of Property/Relation*

There are no genuinely emergent properties in the MH. Rather, there are structurally supervenient properties. And when structural supervenience is applied to the MH, it becomes a form of mereological supervenience.[18]

---

[17] A crucial resource for comparing classic Aristotelian substances and aggregates is Brown, C. M. (2005). *Aquinas and the ship of Theseus*. London: Continuum. See especially chapter 6.

[18] Emergence and structural supervenience were defined in Chapter 2.

> *Mereological Supervenience*: If some property P at some higher level is mereologically supervenient on entities at the next lower level, then P is a new structural arrangement composed of the particulars and their properties at that next lower level.

For example, at the level of molecules, H_2O mereologically supervenes on two atoms of hydrogen and one atom of oxygen placed in the right relational structure. Insightfully, Andrew Bailey calls structural, including mereological supervenient properties, "group-level" properties, and says: "It makes some intuitive sense to continue thinking of group-level properties as being at a higher level than the properties of their components, but, in fact, they are really just very complex concatenations of so-called bottom-level properties."[19]

Strictly speaking, supervenience is a mere doctrine of property (or global) covariance.[20] But most philosophers pack into it different versions of dependence, depending on whether the supervenience relation is taken to be realization, causal, or grounding in genuine emergence, or composition. Mereological supervenience involves the latter and, thus, the ontological dependence it expresses is part-priority—wholes grounded in/dependent upon the separable parts that compose them. Emergent properties or new properties of new substances are either simple (smooth, qualitatively homogeneous properties that are neither structural nor contain particulars as proper parts, e.g., being painful) or exhibit *property-complexity* (given simple properties P and R, the property Q exhibits property-complexity *iff* Q=P+R).[21]

Further, the relations among the separable parts that constitute MAs and structural/mereological supervenient properties are external relations, which we stated in Chapter 3 as,

> *External Relation*: R is an external relation between *a* and *b iff* (i) facts about R are not grounded in facts about the natures of *a* and *b*; and (ii) possibly R fails to hold between *a* and *b*, while at least *a* or *b* remains unaltered.

If the relation R of *a*R*b* is external to *a*, then *a* can exist and retain its identity if ¬ *a*R*b*.

7.2.4 The Structure of the Mereological Hierarchy

There are four general features of the MH's structure that are important for our purposes. First, the direction of the MH levels. As one of us has showed elsewhere, while most talk about the hierarchy going up from bottom to top, this is wrong.[22] The mereological

---

[19] Bailey, A. (1998). The five kinds of levels of description. In S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.), *Towards a science of consciousness II*: *The second tucson discussions and debates* (pp. 577–583). Cambridge, MA: MIT Press. The quote is from 581.

[20] See Kim, *Mind in a physical world*, chapter 1.

[21] Our view is slightly different from Armstrong's since (1) we don't believe it possible that all properties are complex, (2) Armstrong's use of & in Q&R, coupled with his label of "conjunctive" properties seems to be a mistake. Logical connective like conjunction ("&") are applicable to propositions, but not to being rational and being an animal. See Armstrong, D. M. (1978). *A theory of universals: II: Universals and scientific* realism (pp. 30–36). Cambridge: Cambridge University Press.

[22] See Moreland, J. P. (2016). Why top-down causation does not provide adequate support for mental causation. In T. M. Crisp, S. L. Porter, & G. A. Ten Elshof (Eds.), *Neuroscience and the soul* (pp. 51–73). Grand Rapids, MI: William B. Eerdmans.

"hierarchy" doesn't go up; it goes out. From micro-physical particles to atoms, molecules, cells, organisms, ecosystems, etc., the movement is toward wider and wider configurational patterns at the same level of reality, toward larger units of interaction. Thus, all the examples of structural supervenience are not examples of genuine top-down causation, but of outside-in causation.

For example, in illustrating "top-down" causation, Helen Steward advances the frequently used molecule-in-a-wheel case.[23] When force is applied to a molecule outside a wheel, it will travel along a certain path, but if it is part of a rolling wheel, its path will be determined in a "top-down" way by the wheel of which it is a part. Now, it should be clear that a wheel structurally supervenes on a collection of subvenient parts, and this example is outside-in, not genuine top-down causation. Thus, when someone talks about top-down causation in contexts other than emergent properties, this should be taken to mean whole-to-part, outside-in causation.

Second, the interlevel relationship and grounding/dependence. The relation of grounding or ontological dependence runs from the base level-up (bottom-up) (or, as we have noted, inside-out, parts to wholes). Individuals and structural mereological properties at level $n+1$ are grounded in/dependent on those at level n. Ultimately, by way of exhaustive decomposition for both categories, grounding and dependency stop at the fundamental level of micro-physics.

The dependence/grounding is of two sorts: essentially$_1$ grounded and rigid existentially grounded:[24]

> *Essentially$_1$ Grounded*: (IG) x is essentially$_1$ grounded in $y =_{\text{def.}}$ there is a two-place predicate "F" such that it is part of the identity of x that x is related by F to y.
>
> *Rigid Existentially Grounded*: (RG) x is rigidly grounded in $y =_{\text{def.}} (\mathrm{E}x \rightarrow \mathrm{E}y)$.

IG goes beyond existence grounding by specifying what is necessary for x to exist as the kind of thing it is. RG specifies that the existence of x depends on a specific entity (y). The nature (if there is such) and existence of higher-level wholes are ontologically dependent on the specific (separable) parts at the adjacent lower level that compose those wholes.

We may also characterize the interlevel dependency through what is called

> *Serial Bonding*: The *x*s compose something *iff* the *x*s are atomic simples bonded in ways R1 or R1$^{/}$ or R2$^{/}$, and so on, or the *x*s are things composed of atomic simples bonded in ways R2 or R2$^{/}$, and so on, or the *x*s are things composed of things composed of atomic simples bonded in ways R3, and so on.

It will be important to keep *Serial Bonding* in mind when we argue below that MAs at all levels above the bottom are subject to ME and turn out to be mereological inconstant objects.

Third, the existence of a top level. Most part-priority advocates of the MH, especially if they reject Unrestricted Composition, accept the existence of a top level beyond which part/whole composition fails to occur. However, a small handful of philosophers reject this due to the possible existence of *Junky Worlds*: Some world *w* is junky *iff* every object in

---

[23] Steward, H. (2012). *A metaphysics for freedom* (pp. 233–243). Oxford: Oxford University Press.

[24] Inman, pp. 62–74.

w is a proper part. Others have argued against this claim, and it is beyond our interests to consider it.[25] We side with those who take there to be a top level.

Fourth, the existence and nature of the bottom level. Most advocates of the MH accept the existence of a fundamental, ungrounded base of atomic simples, usually

> *Priority Microphysicalism*: The microphysical separable parts of composite wholes are fundamental simples (without proper parts), usually substances or some other units (e.g., a unit of wave-function stuff), and are metaphysically prior to their wholes.

Andreas Hutteman unpacks priority microphysicalism into three theses that taken together comprise it:[26]

> *Micro-Determinism*: The behavior or the properties of compound systems are determined by the behavior or the properties of their constituents and the relations among them but not vice versa.
>
> *Micro-Government*: The laws of the micro-level govern the systems on the macro-levels.
>
> *Micro-Causation*: All causation takes place in virtue of the causation on the level of the (ultimate parts)—or the microphysical level. Macro-causation is entirely derivative and piggybacks on the causation of the micro-constituents.

There are those who reject the existence of a bottom level, usually by embracing the existence of *atomless gunk*: material objects each of whose parts have proper parts. Some material object *x* is *mereologically* gunky just in case *x* has no atomic parts. Finally, a material object *x* is a *gunky body iff x* has proper parts but no ultimate proper parts. On a gunky view, the levels continue to descent infinitely, never reaching a bedrock, fundamental base. Thus, advocates of *atomless gunk* reject the notion of part-priority because they do not accept the existence of fundamental parts upon which higher-level wholes depend. We will examine the notion of *atomless gunk* and stuff/masses later in this chapter. For now, we assume there is a fundamental level.

Closely related to *atomless gunk* is the notion of

> *Homogeneous Continua*: The *x*s compose something *iff* they form a homogeneous continuum (a material thing with no sharp internal boundaries.)[27]

Koons and Pickavance point out that homogeneous continua do not have actual parts, only potential ones and, thus, material homogeneous continua turn out to be

> *Aristotelian Atoms*: *x* is an *Aristotelian atom iff x* is an atom (with no actual proper parts) and *x* has the passive power to bring into existence one or more fission products.

Composition does not apply to homogeneous continua since they have no actual proper parts. And while it seems that the base level of the hierarchy could consist in units of homogeneous continua, this is not a popular view among the advocates of the MH who are our focus. So, we set it aside.

---

[25] Inman, pp. 203–207.

[26] Hutteman, A. (2004). *What's wrong with microphysicalism* (pp. 7–8). London: Routledge.

[27] Koons & Pickavance, p. 511.

7.2.5 Causation

Advocates of the MH accept some favored standard account of efficient causation (e.g., Humean, counterfactual, law-governed, dispositional powers) and laws of nature, along with one of the many options for depicting the relationship between causation and laws. In the strict, philosophical sense, such causation is limited to event-event causation at the bottom level. Irreducible teleology is rejected. In Chapter 9, we will see that staunch libertarian agency, especially agent-causation, is highly implausible, given MH.

What about levels higher than (wider than) the bottom one? This raises the issue of so-called "top-down" causation. Regarding the precise function of "top-down" causality, there is a view according to which wholes exert efficient causality on their parts by intervening into the chain of subvenient events and changing the direction of that chain. In this sense, top-down efficient causation involves causal activities that "push and pull the lower level components of the system."[28]

As far as we can tell, this view is universally rejected by the advocates of top-down causation we are seeking to criticize. Nancey Murphy argues that such a notion of "top-down" causation may require the postulation of *spooky new entities*, it threatens the integrity of the special sciences, and it smacks of interventionism according to which wholes intervene in the course of micro-physical processes and, as it were, interrupt them.[29] Helen Steward opines that this notion of causation is "*spooky, magical, dualistic*, and otherwise metaphysically problematic …"[30]

In light of these remarks by Murphy, Steward, and many others, an alternative view of "top-down" causation is offered: "Top-down" causation for structural supervenient properties should be cashed out, not in terms of efficient causality, but in terms of causal constraints. Various distributions of a whole's parts are consistent with micro-physical laws, but that one distribution was selected as opposed to another must be explained by referring to "higher-level" entities. Thus, "top-down" causation functions as "a larger system of causal factors having a selective effect on lower-level entities and processes."[31] Further, higher-level entities (e.g., wholes, processes, relational configurations) provide context-sensitive constraints, e.g., boundary conditions, feedback loops, nonlinear factors, on the parts at "lower" levels, and higher-level entities are what sustain the sequence of synchronous arrangements of subvenient parts (e.g., reference to a whirlpool and its "top-down" causal influence must be made if we are to explain the relatively stable sequencing through time of subvenient parts arranged whirlpooly).

7.2.6 Concepts and Laws

It is almost universally acknowledged by advocates of the MH that there is no "top-down" intervention or interference with the basic laws of physics and the flow of events they govern at the bottom level. But those advocates embrace the existence of emergent laws, concepts, and theories at higher levels. Let's examine how to understand this.

---

[28] Clayton, P. (2004). *Mind and emergence* (p. 24). Oxford: Oxford University Press.

[29] Murphy, N. (2007). Reductionism: How did we fall into it and can we emerge from it? In N. Murphy & W. R. Stoeger (Eds.), *Evolution and emergence* (pp. 19–39). Oxford: Oxford University Press.

[30] Steward, pp. 228.

[31] Murphy, p. 29. Cf. Steward, pp. 232–245.

First, there are what William Hasker calls *emergence*$_{1a}$.[32] This involves new structural properties that leave the base of microphysics alone and involve no novelty or top-down causation. From a God's eye perspective, all the behavior of the whole can be predicted and explained entirely by those of the proper parts. However, Murphy adopts a realist approach to such higher-level properties, concepts, and laws in a way that defeats causal reduction. She claims that this sort of emergence provides a new property that gives rise to a new law governing the whole's behavior when environmental influences are taken into account. Thus, emergent properties and laws undercut causal reduction.[33] As an example, Murphy points to a paper airplane. The plane's shape is a holistic emergent property that, together with environmental inputs (the force of the hand that throws it, air current), produces emergent behavior and, in principle, a new law.

To see her point, consider Newton's First Law of Uniform Linear Motion: Objects that do not experience any net force will continue to move in a straight line at constant velocity. Thus, if a sheet of metal is perfectly flat and air moves past the sheet with no change in net forces and absent friction/resistance, the air will move at constant velocity in a straight line. However, if the sheet's first half, say the first two feet, remain the same, but the second half is bent in a 30° angle, airflow will be changed and, in principle, a new (at least Humean) law could be developed to describe and predict airflow past the bent sheet.

But is this really an example of top-down causation and a defeater for causal reduction? We agree with William Hasker who rejects Murphy's case and her understanding of *emergence*$_{1a}$.[34] Hasker points out that the "new" property is a straightforward consequence of the configurational arrangement of the plane's elementary particles. There is, indeed, a different law of the whole's motion given this structure and environmental interaction that does not apply to the plane's parts taken in isolation. But this point, along with the laws of aerodynamics in no way avoid causal reduction.

Setting aside a straightforward realist understanding of higher-level concepts, laws, and causation, a second approach to these matters comes from staunch naturalist and strict physicalist David Papineau.[35] Let us call all levels above the base *Special Science levels*. All sciences except microphysics are special sciences. For our purposes, Papineau holds that there are two basic types of special sciences: those with teleological underpinnings and those without such. The latter are type-type reducible to physics and, thus, the concepts they employ carve the world at the joints, albeit in a broader way than those of physics. Chemistry is the paradigm case of such a science. For example, the concepts of the elements of the periodic table—e.g., gold—are type reducible to the concepts of physics—e.g., having atomic number 79. The laws and causal features of these sciences are real and, in principle, type-reducible to those of physics.

The former are sciences like biology, psychology, and sociology. They employ functional concepts characterized in terms of two components: (a) a concept type consisting of input-realizer-output (I→R→O); (b) variable physical tokens that serve as realizers. The problem here is to explain how it is that there can be lawlike relations between I and

---

[32] William Hasker, pp. 173–174. For a taxonomy of emergence, see chapter 2, section 2.2.4.

[33] Murphy, N. (2005). Nonreductive physicalism. In J. B. Green and S. L. Palmer (Eds.). (pp. 115–138). Downers Grove, IL: InterVarsity Press. The example and relevant discussion come from 135–138.

[34] Hasker, W. An emergent dualism response. In Green and Palmer, pp. 143–146.

[35] Papineau, D. (1993). *Philosophical* naturalism (pp. 33–51). Oxford: Blackwell.

O with a large number of physically different intermediate realizers. Papineau's answer is to adopt an etiological view of teleology for functions and the like. He goes on to argue that due to evolutionary selection, the various realizers produce the same (or similar) output, usually given similar inputs, because there were selection mechanisms at work that "chose" those realizers precisely because they produced the same output.

So, concepts in these special sciences are to be understood via anti-realism and functionalism, they do not carve the world at the joints, and the laws of these sciences are adopted for their empirical adequacy. No new causal powers are involved. In these sciences, there is no higher-level causation, their concepts are useful fictions, and empirical/predictive accuracy is what we expect from their laws. These special sciences avoid reduction by advancing irreducible concepts/laws that are epiphenomenal and anti-realist.[36]

Jaegwon Kim makes an important distinction between levels and orders in the MH to avoid what is called the

> *Generalization Argument*: If there is a problem with mental causation at the mental level, there should be the same problem with every level in relation to lower levels. If there isn't such a problem at these other levels, then why think there is one at the mental level?[37]

Kim's answer is to distinguish levels from orders. On the MH, levels are interlevel distinctions constituted by the part–whole relation and a realist view of entities (atoms, molecules, rocks) at higher levels. Orders are intralevel distinctions constituted by the realization relation and a functionalization of the "properties" involved. These functional properties are actually just concepts, ways of taking, say, a brain state. In this way, the problem of mental causation does not generalize since it involves a progression of orders and not levels. As a result, special sciences that are functionalized in this way employ epiphenomenal concepts.

7.3 A DEFENSE OF PREMISE (4): HUMAN PERSONS ARE ENDURING SIMPLE SPIRITUAL SUBSTANCES

7.3.1 Simple and Complex Views of Diachronic Personal Identity

In the current literature, there are two main views of personal identity (PI), with various versions of each: the simple and complex view.[38] The simple view can be understood in three mutually inclusive ways: (i) PI is absolute, conforming to the indiscernibility of identicals. (ii) PI is grounded in a mereologically simple (has no separable parts) entity, the person. (iii) PI is type and token (metaphysically) irreducible to some analysis, usually construed in terms of necessary and sufficient psychological or bodily criteria, in which "person" drops out.

Likewise, the complex view can be understood in three mutually inclusive ways: (i) PI is not about sustaining strict identity; rather, it is about a weaker relation of persistence spelled out in various ways. (ii) PI is analyzable into a set of necessary and sufficient conditions that involve psychological states or bodily criteria (which sometimes seeks to embrace organisms as animals that are genuine enduring continuants) and that paraphrase away "person" or "personal." (iii) The entity that grounds personal persistence is mereologically complex and is constituted by temporal parts related in the right way by some relation(s) R, or by successive, nonidentical MAs or bundles of properties, or in some other way.

---

[36] For a very insightful discussion of these matters, see Robinson, H. (2016). *From the knowledge argument to mental substance* (pp. 156–158). Cambridge: Cambridge University Press.

[37] Kim, pp. 79–87.

[38] See Gasser and Stefan, pp. 1–17.

We will discuss complex views below when we analyze and evaluate the persistence of MAs. At this point, we proffer arguments for the simple view.

7.3.2 Our Basic Awareness of Ourselves

There are certain things about ourselves that we know and about which we have a right to be sure. When we say we have a right to be sure of them, we do not mean that they are indefeasible or that we should terminate inquiry about them. Rather we mean that our knowledge of these things confers this right and also the knowledge that we do, in fact, know these things. We, thus, have the intellectual right to rely on this knowledge in assessing other issues, and to employ things entailed by this knowledge in developing our views of the world and so forth.[39] These claims were defended in Chapter 4.

Thus, Stewart Goetz has advanced an argument according to which we are directly aware of ourselves as indivisible, simple, spiritual substances.[40] We have modified the argument to fit with diachronic personal identity. Suppose you are approaching a brown table and in three different moments of introspection you attend to your own awarenesses or experiences of the table. At time t_1 you are five feet from the table, and you experience a slight pain in your foot (P_1), a certain sensation of a light brown table from a specific place in the room (S_1), and a specific thought that the table seems old (T_1). A moment later at t_2 when you are three feet from the table you experience a feeling of warmth (F_1) from a heater, a different table sensation (S_2) involving a different shape and slightly different shade of brown than that of S_1, and a new thought that the table reminds you of your childhood desk (T_2). Finally, a few seconds later (T_3) you feel a desire to have the table (D_1), a new table sensation from one foot away (S_3), and a new thought that you could buy it for less than twenty-five dollars (T_3).

In this series of experiences, you are aware of different things at different moments. However, at each moment of time, you are also aware that there is a self at that time that is having those experiences and that unites them into one field of consciousness. Moreover, you are also aware that the very same self that had the experiences at T_1, T_2, and T_3. Finally, you are aware that the self that had all the experiences is none other than you yourself. This is pictured in Figure 7.2:

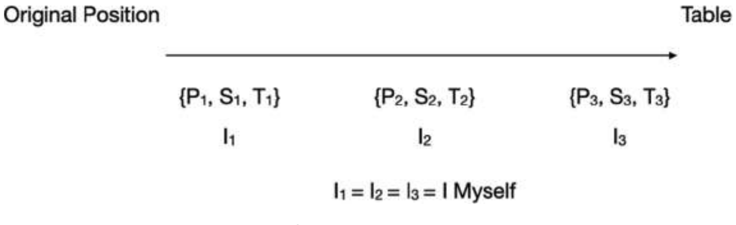


FIGURE 7.2 Temporal Self-Awareness.

---

[39] See Chisholm, R. M. (1977). *Theory of knowledge* (2nd ed., pp. 116–118). Englewood Cliffs, NJ: Prentice-Hall.

[40] Goetz, S. (2001). Modal dualism: A critique. In K. Corcoran (Ed.), *Soul, body and survival* (p. 89). Ithaca, NY: Cornell University Press.

Through introspection, you are aware that you are the self that owns and unifies your experiences at each moment of time and that you are the same self that endures through time. This is obvious to most people. When one hums a tune, one is simply aware of being the enduring subject that continues to exist during the process. This is a basic datum of experience, and it grounds the properly basic belief that one is a simple immaterial enduring substance. One also becomes directly and positively aware (some prefer strongly conceives) of one's self as simple and essentially immaterial (i.e., intrinsically constituted by categorial and dispositional properties/powers of consciousness.)

Goetz argues that this direct awareness of a simple spiritual self is what epistemically grounds one's awareness that one is distinct from and contingently related to one's body. The question of contingency will be taken up in Chapter 8, but the distinctness is based on the awareness that one is not composed, is essentially/intrinsically conscious and the knowledge that one's body is composed and physical.[41]

If successful, this argument shows that SD is not simply a theory postulated to explain certain data (e.g., a proper ground for libertarian freedom). It is also a report of that of which one is directly aware. There are three major objections to this argument.

First, some argue that third-person knowledge, especially supported by neuroscience, trumps alleged first-person knowledge, so the argument is defeated. We will address this problem in Chapter 10.

Second, Peter van Inwagen responds to the argument by claiming "… when I enter most deeply into that which I call *myself*, I *seem* to discover that I am a living animal. And, therefore, dualism seems to me to be an unnecessarily complicated theory about my nature."[42]

Clearly, we have a case of dueling intuitions. But there is more to say than this. In Chapter 5 we offered data from studies in cognitive science to the contrary and offered replies to those such as van Inwagen's. In the Appendix we offer a more detailed presentation and defense of these studies.

Third, Paul Churchland and many others claim that there is no such thing as knowledge by acquaintance, direct awareness, or simple seeing. According to Churchland, all seeing (awareness) is seeing as or seeing that, i.e., constituted by concepts or attitudes toward propositions.[43] Consequently, even if we have selves, we do not have direct awareness of them.

We offer two responses to this argument. For one thing, below and in Chapter 8, we will explain what Edmund Husserl called "fulfillment structures." These are cases where someone begins with a fairly empty, lean concept of something (a specific book is ready for pickup in the bookstore), and one lives through a series of steps until the object itself (the specific book) is directly present and the original concept has been filled out and verified. These cases are experienced daily by us all in which we verify or correct an initial concept by “stepping outside of the concept,” having direct awareness of the concept’s intentional object, and comparing the concept with the object itself given directly to our awareness.[44] In Chapter 9, we will examine how fulfillment structures go beyond sense perceptible objects and include direct awareness of abstract properties/relations and of one’s self.

---

[41] In Chapter 8, we will examine and reject contingent physicalism.

[42] van Inwagen, P. (1995). "Dualism and materialism: Athens and Jerusalem?" *Faith and Philosophy*, *12*(4), 475–488. The quote comes from p. 476.

[43] Churchland, P. (2013). *Matter and consciousness* (3rd rev. and expanded ed., pp. 130–131). Cambridge, MA: MIT Press.

Further, recall from Chapter 5 our argument from what we called the

Causal-Acquaintance Principle: (CA) $(s)(x)(y) \rightarrow (K_a sxK_a sy)$[45]

CA says that necessarily, if a subject *s* has knowledge by acquaintance with a causal fact *x*, then *s* has knowledge by acquaintance with the relevant causal object *y*. As we argued, the difference between active and passive thoughts shows that one has knowledge by acquaintance with one’s self.

7.3.3 Plantinga’s Replacement Argument

Alvin Plantinga has presented a relevant argument that we find to be quite persuasive.[46] Plantinga employs a replacement argument to demonstrate that it is broadly logically possible for one to exist when one’s body does not, though he does not explicitly argue that one could exist disembodied with no body at all (though he believes that to be the case). Putting his argument in the first-person as he does, according to Plantinga, I am a substance; if I am a physical substance, I would be my body, or a proper part of it or an object distinct from my body but composed of the same matter and collocated with it. But I am none of these things. Call any of these relevant objects “B.” It is possible that I continue to exist when B doesn’t, so I am not identical to B.

At the macrolevel, science could advance such that I continue to read the paper throughout a process during which each macropart of my body is replaced by others in a microsecond and the original parts are annihilated. If the process occurs rapidly, my body will not continue to exist, but I do since I myself continue to read the paper. Plantinga advances the same argument for the brain. And the same goes for a material object constituted by the same matter as and collocated with my body. It, too, would be annihilated if God destroyed all of its parts, whether micro- (e.g., cells) or macrolevel parts.

But, asks Plantinga, would the replacement really kill B? After all, all the matter of the body is replaced regularly, and the body seems to survive. Why should merely shortening the time interval make a difference? Plantinga’s answer is that speed kills. It takes time for new parts to be assimilated into the bodily whole, so he makes the replacement time quicker than the assimilation time.

Peter van Inwagen has responded to Plantinga’s replacement argument.[47] According to van Inwagen, the replacement scenario isn’t possible because I am identical to my body.

---

[44] For more on this, see Smith, R. S. (2012). A positive case for our knowledge of reality. In R. S. Smith (Ed.), *Naturalism and our knowledge of reality* (pp. 182–196). New York: Routledge.

[45] Where *s* ranges over knowing subjects, *x* ranges over causal facts, for example, a hammer’s causing a nail to move, and *y* ranges over the associated causal objects that constitute their causal facts (for example, the hammer). K_a is “has knowledge by acquaintance with.”

[46] Plantinga, A. (2007). Materialism and Christian belief. In P. van Inwagen & D. Zimmerman (Eds.), *Persons: Human and divine* (pp. 102–105). Oxford: Clarendon.

[47] van Inwagen, P. (2007). Plantinga’s replacement argument. In D.-P. Baker (Ed.), *Alvin Plantinga* (p. 188ff). Cambridge: Cambridge University Press.

To keep this from being a question-begging assertion, van Inwagen invites reflection on this question: Is it more plausible to suppose that it [Plantinga's scenario] is possible than it is to suppose that I am identical with my body? We believe the answer to this is not only "yes," but clearly so. The dualist intuitions shared worldwide by most people, including children, seem to be quite reasonable to most folks. Moreover, when most people introspect, they do not confront their "composed animality," as van Inwagen has suggested, but their own unified, simple conscious self and its states.

Van Inwagen agrees that if the replacement of parts in Plantinga's envisioned thought experiment is more rapid than the assimilation time, then B is destroyed. But why think, he queries, that I continue to exist throughout the process? For Van Inwagen, Plantinga seems to argue for this on the basis of continuous consciousness, i.e., that a single, unified process of consciousness exists throughout the entire replacement process and a necessary condition for this is that a single person must be the enduring conscious subject of that process. And it seems to be a fundamental datum of self-awareness that I am an enduring continuant throughout processes like reading the paper. Nothing, it seems, is more evident than the fact that as I experience a series of notes in a tune, I am the unifying I for the extended experience.

Moreover, Plantinga has responded to van Inwagen's argument in the following way:[48] Let t and t* be the beginning and end of the replacement scenario, respectively. According to Plantinga, van Inwagen thinks his argument is this one:

(1) The Replacement Scenario is possible.
(2) If the Replacement Scenario were actual, I would exist at t, during all of t to t*, and at t*.
(3) If the Replacement Scenario were actual, B, my body, would not exist at t*.
(4) Therefore, possibly, I exist when B does not and I am not identical to B.

Van Inwagen attacks (2), but Plantinga claims that (2) isn't part of his argument and van Inwagen has misunderstood his case. Rather, the case goes like this:

(1*) Possible, the Replacement Scenario is actual, and I exist and am conscious at t, from t to t*, and at t*.
(3) If the Replacement Scenario were actual, B, my body, would not exist at t*.
(4) Therefore, possibly, I exist when B does not, and I am not identical to B.

Here, (2) is not asserted or presupposed. Accordingly, Plantinga sees no reason to reject the modal intuition that lies under the Replacement Argument. We leave to the reader to evaluate this dialectic, though our sympathies clearly lie with Plantinga. It would be open to a Plantingian critic to argue that we are atomic simples that can endure the sort of part replacement in the thought experiment. We will look at this response below.

But, perhaps, there just is no single person who endures this process, or maybe there is not a single conscious episode, but merely a series of discrete conscious states. We have already given defeaters for these claims, but we turn to an additional consideration relevant to these assertions.

---

[48] Plantinga, A. (2007, Fall). Reply to Peter. Paper presented to the Department of Philosophy, Notre Dame University. Permission granted by Alvin Plantinga to cite from this unpublished paper.

7.3.4 Normative, Teleological Rationality, Rational Deliberation, and the Simple Enduring Spiritual Self

We agree with those epistemologists that epistemology is prescriptive and normative and cannot be reduced to or replaced by descriptive naturalized epistemology. Whether understood deontologically (there are intrinsically correct epistemic rules and imperatives it is one's rational duty to follow) or in terms of obtaining in the right way (e.g., according to virtue epistemology, intrinsically good justified or warranted mental states), rationality (rational justification or warrant) is intrinsically normative.

That said, there is another aspect to such rationality that manifests itself in at least two forms of reasoning: rational deliberation and fulfillment structures, and these manifestations seem to entail a view of the self as an enduring simple spiritual (essentially characterized by certain properties/states of consciousness) substance:

7.3.4.1 Reasoning Is Intrinsically Teleological

Processes of reasoning are intrinsically teleological in that the precisely ordered sequence of states constituting such processes are done *for the sake of* some end, namely, achieving truth or the best overall epistemic state. The sort of real, intrinsic teleology advanced by Aristotle and his followers and constituting rational processes entails an enduring substantial continuant of the right and relevant kind, apt for unifying and guiding the process. From early to final stages, the teleologically ordered series of lawlike events occurring for the sake of their proper end all take place within the same entity that endures over time. Such a series is an inherent *movement* of steps toward the end.

As William Hasker has correctly noted, matter as conceived by science since the 1700s does not contain/exhibit intrinsic, irreducible teleology.[49] Thus, the following argument seems sound:

(1a) If strict physicalism is true, there is no irreducible, intrinsic, rational teleology.
(2a) Rational deliberation exhibits irreducible, intrinsic, rational teleology.
(3a) There is such a thing as rational deliberation.
(3b) Therefore, strict physicalism is false.
(4a) If there is irreducible, intrinsic, rational teleology, then there is an enduring, simple, rational, substantial, deliberative self.
(5a) There is an enduring, simple, rational, substantial, deliberative self.

Reasons explanations are irreducibly teleological (Frank raised his arm because he voted or had a desire to vote), and various inductive or deductive thought processes reason through a series of steps *in order to* or *for the sake of* reaching a sound, true, rational conclusion.[50] Besides the analysis of reasons explanations, when one attends to one's own endeavorings, it becomes introspectively evident that the various steps in such processes are formulated for the sake of drawing a particular conclusion. If one pays attention to fairly simple mental states in second-order awarenesses of first-order states, then it seems reasonable to say that unless there are substantial, non-question-begging, overriding defeaters, then one should believe that things are as they seem. For example, a pain is as it seems to be in such acts. Similarly, one's own teleological endeavorings are as they seem.

---

[49] See William Hasker, pp. 64–109. See also, Hoffman, J., & Rosenkrantz, G. S. (Eds.). (1997). *Substance: Its nature and existence* (pp. 98–99). London: Routledge.

[50] See, e.g., Rickabaugh, B., & Buras, T. (2017). The argument from reason, and mental causal drainage: A reply to van Inwagen. *Philosophia Christi*, *19*(2), 381–399.

Additionally, when one attends to the different states containing propositional contents in rational sequences that constitute the inductive or deductive premises of the sequence, it becomes evident that these states are *means—rational means*—to the end of drawing the conclusion. And when one attends to both the drawing of the conclusion and the conclusion so drawn, it becomes evident that the conclusion is the *end* for the sake of which the process was undergone. In fact, if one is reasoning to a conclusion, and one line of reasoning is seen to be (or not to be!) epistemically useful for drawing a conclusion, through introspection we are able to become aware of this and thereupon adopt an alternative line of reasoning as a subsequent means to the conclusive end. In this way, the teleological view of deliberation supports the derivation of compensatory counterfactuals. Consider the following:

> (T) S E'd (reasoned through a deliberative series) in order to C (draw a conclusion about what insurance policy to purchase).
> (T_C) If E'ing had required G'ing (gathering such and such information), then S would have G'd.

An irreducibly teleological interpretation of T provides a powerful, straightforward way of deriving T_C. Moreover, the teleological end is essential to the sequences in the chain (E'ing) being what they are—precisely ordered means to an end. Had S not been directing S's rational process toward C'ing, S would not have E'd.[51]

Reductive analyses of these teleological facts about deliberation cannot appeal to staunch naturalism or strict physicalism on pain of begging the question, and we know of no good argument that provides a sufficiently strong over-riding defeater for the types of evidence for teleology just cited and of which we are aware daily in our thought processes.

That the employment of such inherent teleological rationality entails an enduring, simple, rational, substantial, deliberative self seems evident in two areas: rational deliberation (already touched on) and fulfillment structures.

Starting with rational deliberation, let's look at thinking through the *modus ponens* syllogism. As A. C. Ewing argues, there must be an identical rational self that endures through the rational process.

> To realize the truth of any proposition or even entertain it as something meaningful the same being must be aware of its different constituents. To be aware of the validity of an argument the same being must entertain premises and conclusion; to compare two things the same being must, at least in memory, be aware of them simultaneously; and since all these processes take some time the continuous existence of literally the same entity is required. In these cases an event which consisted in the contemplating of A followed by another event which consisted

---

[51] For more on these points, see Sehon, S. R. (1994). Teleology and the nature of mental states. *American Philosophical Quarterly, 31*(January), 63–72.

> in the contemplating of B is not sufficient. They must be events of contemplating that occur in the same being. If one being thought of wolves, another of eating, and another of lambs, it certainly would not mean that anybody contemplated the proposition ‘wolves eat lambs’ … There must surely be a single being persisting through the process to grasp a proposition or inference as a whole.[52]

If the conclusion of a syllogism is to be grasped *as a conclusion*, it must be *drawn* from the experiences of each premise singularly and then, together. As Ewing notes, a successive series of I-stages or a person constituted by temporal parts as in worm theory cannot engage in such acts; only an enduring I can. Moreover, if the rational agent who embraces the conclusion is to be regarded as intellectually responsible for his reasoning, it must be the same self at the end of the process as the self who lived through the stages of reasoning that led to drawing the conclusion. One is not responsible for the acts of others or of other person-stages. So intellectual responsibility seems to presuppose an enduring I. This I is a substance constituted by (at least) consciousness.

Various chapters in this book provide additional defenses of a substantial, simple, enduring mental self, and explain why strict physicalism cannot countenance such a self. But for now, we simply note the basic reason for this inability: If one starts with separable parts, and simply re-arranges them according to natural laws into new relational structures constituted by external relations, then in the category of individual, one’s ontology will have atomic simples (*sans* difficulties with such) and MAs with insufficient unity to avoid ME or to ground absolute unity at and through change of parts.

7.3.4.2 Fulfillment Structures and Teleologically Arranged Action Plans

Further, there is the matter of fulfillment structures (first discussed by Husserl; see Chapter 5.4.1) and first-person epistemic success. We are capable of engaging in first-person examination of our own mental—e.g., conceptual and propositional—contents, engaging in a series of mental steps that would be relevant to verifying or falsifying those contents, eventually becoming acquainted with the thing itself (the intentional object of the original mental contents), and comparing our initial mental contents with the thing itself in order to experience the truth (or falsity) of those contents.

During such “fulfillment structures,” one may begin with fairly empty (little content) concepts. Reflection on them leads to series of conscious states grasped precisely *as* a succession in which those concepts are filled more and more until fulfillment is reached—the thing itself is present to consciousness. Fulfillment structures are teleologically arranged action plans leading to (hopefully) fulfillment. The same enduring conscious/rational self formulates, guides, and is present at the end of such action plans as an epistemically responsible subject. These are cases of first-person epistemic success.

David Lund summarized our argument(s) and makes an additional, important point:

> The next major stage in the overall case for acknowledging the reality of a subject maintaining a “strict” identity through time as an immaterial individual consists of an attempt to show that very subject of which one has knowledge by acquaintance

---

[52] Ewing, A. C. (1973). *Value and reality* (p. 84). London: George Allen and Unwin. For a slightly different argument, see Rosenkrantz, G. (2005). An epistemic argument for enduring human persons. *Philosophy and Phenomenological Research*, *LXX*(January), 209–224.

> and whose singular viewpoint is needed for any plausible account of diachronic unity. Diachronic unity rests upon a diachronic singularity made possible by the persistence of the very same subject that provides diachronic singularity.[53]

Elsewhere, Lund continues:

> Thinking through a complex argument is a continuous mental process that requires sustained mental effort and control of attention throughout its duration. It is a process in which one's role as an agent seems impossible to doubt—an agent who initiates the process and then continually wills, directs, controls and sustains it. The continuity of the process seems entirely dependent upon the continuity of agency, which, in turn, renders irresistible the conclusion that the sameself agent persists throughout the process. Thus the singularity of an agent willing and sustaining the process seems as incontestable as the singularity of the consciousness that is able to grasp the process as a succession of events through which it has persisted.[54]

Lund's reference to "the singularity of the agent" means the simplicity of the agent (no proper parts). On his view, synchronic or diachronic consciousness is holistic and not atomistic, and this fact is grounded in the holistic, non-atomistic agent/subject. The most reasonable ground for this sort of unity of consciousness is a simple soul/self (see Chapters 4, 5, and 6).[55]

Considering only the arguments in this section, we think that our conclusion—human persons are enduring simple spiritual substances—is, in Chisholm's term of epistemic appraisal—beyond reasonable doubt: P is beyond reasonable doubt for $S=_{\text{def.}}$ S is more justified in believing p than in withholding p.[56]

7.4 DEFENSE OF PREMISE (1): MAS ARE MEREOLOGICAL INCONSTANT OBJECTS

Unfortunately, our appraisal may be demoted when we consider defeaters and alternatives to treating human persons, their bodies, and brains, as MAs. We have already shown that staunch naturalists and strict physicalists should depict human persons in such a way, perhaps, by reducing them to their bodies or brains, and employing a version of the Complex View to spell out in what way such entities persist. Accordingly, in this section, we will examine this standpoint. In Section 7.5, we will evaluate two physicalist alternatives to the one considered here: we are atomic simples or organisms construed as dynamic biological systems.

---

[53] Lund, D. (2014). Materialism, dualism, and the conscious self. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 56–78). London: Routledge, especially 72.

[54] Ibid., p. 73.

[55] Cf. Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press; Moreland, J. P. Substance dualism and the unity of consciousness. In *The Blackwell companion to substance dualism* (pp. 184–207). For a similar argument for a simple, substantial mental subject based on counterfactuals of origin and not temporal endurance, see Robinson, H. *From the knowledge argument to mental substance* (pp. 233–247).

[56] Chisholm. (1989). *Theory of knowledge* (3rd ed., p. 11). Englewood Cliffs, NJ: Prentice-Hall.

7.4.1 An Ontological Assay of MAs

Why should we believe ordinary material objects construed as MAs mereologically inconstant? The short answer is: *Given part alteration, strict physicalism does not have the ontological resources to ground or account for the type of unity needed for an MA to be a continuant*. It may even be that physicalism does not have the ontological resources to avoid eliminativism regarding MAs, but that issue must be left for another time.

A proper metaphysical analysis of MAs provides no entity adequate to ground metaphysically their literal identity through part alteration. To illustrate, suppose we have some MA W, say a car, in the actual world w at some time t, and let "the *p*s" refer distributively to all and only the atomic simples (assuming such) that make up W. Given that the *p*s just are a specific list of simples taken distributively regardless of structure, a different list of simples, the *q*s, would not be identical to the *p*s, even if both lists shared all but one part in common. This same insight would be true if we took "the *p*s" and "the *q*s" collectively as referring to some sort of mereological sum. In either case, no entity "over and above" the parts exist to ground sameness through part alteration. Nor can it rationally be asserted that such a plurality-in-flux retains strict identity as a brute fact. The identity of such a many is derivative, not fundamental, and resides in the identity of its parts (and relation-instances between and among them).

W has different persistence conditions from—and, thus, is not identical to—the *p*s. W could be destroyed and the *p*s (taken in either sense) could exist. Now, let S stand for all and only the various relations standing between and among the *p*s. S is W's *type* of structure. Is W identical to S and the *p*s? No. W has its own structure, say in comparison to some other whole W* exactly similar in structure to W. W and W* have their own token structures of the same type. As mentioned in the previous chapter, animalism is committed to something near enough to this, where W's structure is a biological process, a life.

Given that S is a universal, it is not sufficient for individuating W's specific structure. For that we need SI, W's structure-instance, W's token of S, and SI will consist of all and only the specific relation-instances that are instantiated between and among the *p*s. Let "the *r*s" stand for all and only the relevant relation-instances that compose SI. We think it is now obvious that SI is an MA composed of the *r*s. If the *r*s undergo a change of relation-instances, it is no longer the same collection of relation-instances. Given that SI is just an MA or, perhaps, a specific ordering of the *r*s, if the *r*s undergo a change of relation-instances, SI ceases to exist; a different structure (perhaps exactly similar to SI) obtains, since no entity serves as a ground for SI's sameness through part replacement. If the animalist posits a life as an entity, it isn't clear how the animalist is not a substance dualist.[57] If W is the *p*s plus SI, it follows that W is subject to mereological-essentialist constraints.

It is interesting to note that several of the late Medieval Aristotelians noted that a shift from classic substances with substantial forms and prime matter to substances depicted as Democritian atoms resulted in several troubling results: a loss of grounding for the synchronic unity of macro-objects, especially organisms, a loss of teleology and, more important for present concerns, an inability to distinguish metaphysically genuine generation/corruption and mere alteration. This inability meant that the former collapsed into mere alteration with the result that genuine accidental or substantial change is replaced by mere instantaneous succession. Everything is alteration. The clear result is that macro-objects are all MAs construed as mereological inconstant objects.[58]

---

[57] See Rickabaugh, B. (work in preparation). Enduring problems for animalism.

Someone might object that we have deeply engrossed commonsense intuitions according to which the ordinary aggregates of everyday life are continuants even in light of part alteration. However, the source of error for these intuitions is easy to spot: They abandon the Scientific Image and are falsely based on the Manifest Image. And as David Barnett notes, from the latter perspective, we see aggregates as simple blobs and are misled into thinking of them as enduring continuants.[59]

7.4.2 Defeaters for the Complex View of PI

If we are MAs, then some sort of Complex View captures our PI. All variants of this view are committed to some form of the following reductive analysis: P_1 at t_1 = P_2 at t_2 *iff* _______. The analysis will be filled in with (i) some variant of Rx_1x_2 according to which R is some relation and x_1, x_2 are some sort of psychological or bodily state, and (ii) additional conditions to patch up the reductive analysis in light of counterexamples.

The two main versions of the Complex View are the psychological and bodily version. The latter is best identified with animalism and, since we treated it in Chapter 6, we dismiss it here and focus on the former. Thought experiments are central in the literature on PI, and while some reject such experiments, we offer a defense of their reliability in association with modal epistemology and skepticism in Chapter 8. For now, we follow the majority in accepting their use. Thus, we agree with Chisholm when he said, "They [thought experiments] seem to me to be worthy of being taken, at least provisionally, as data in our philosophical enquiries about the person; in other words, we should affirm them until we have very good reason for rejecting them."[60]

We cannot treat all the variants of the psychological version, but the following should be adequate. Complex Views do not preserve absolute PI; they settle for a weaker relation R between person stages adequate for persistence: similarity, chains of stages that are continuous and are (1) non-gappy (some accept gaps) and (2) such that for each pair of adjacent stages, the degree of similarity must be high, along with additional conditions. The states themselves are some combination of memories, character, and personality, values, and goals.

Three major objections are usually raised against psychological views as stated above.

7.4.2.1 The Circularity Problem

The continuity of the relevant psychological factors—let's take memories as an easy example—must be the continuity of *my* memories to be adequate. Which memories are mine? Those that belong to me, a specific person. In response, one can provide a reductive analysis such that *my* and its first-person surrogates disappear and are replaced with some entity described from the third-person perspective. For example, past memories are mine *iff* they stand in the right causal chain to a relevant, current conscious state. But two problems arise with such solution.

---

[58] Pasnau, pp. 552–557.

[59] Barnett, D. (2010). You are simple. In R. Koons & G. Bealer (Eds.), *The waning of materialism* (pp. 161–174). New York, NY: Oxford University Press, especially 170–171.

[60] Chisholm, R. M. (1970). Reply to Strawson's comments. In H. E. Diefer & Milton K. Munitz (Eds.), *Language, belief and metaphysics* (pp. 187–189). Albany, NY: State University of New York Press. Quote is on page 188.

First, it is hard to specify "right" and "relevant, current" so as to defeat counterexamples that show a given analysis is neither necessary nor sufficient, without being implicitly guided by the concept of the first-person *I*, *my*, and so forth. The pervasive, ineliminable need for such concepts in the dialect provides reasons for thinking that they are fundamental, not derivative, as the Simple View would have it.

Second, for every third-person analysis of the right causal chain and the relevant, current conscious state, there is always an unanswered question: But is the current conscious state produced by the right causal chain really mine? This question is no mere explanatory gap or expression of two ways of conceiving/referring to the same thing. Instead, as Richard Swinburne has shown, there is a fact of the matter left out by the third-person reduction. Moreover, the *de se* employment of *I*, *my*, *mine* involves *de re* and not *de dicto* reference.[61] So this problem is no mere Moorean open-question argument to be solved by appealing to two semantically different, co-referring expressions.

7.4.2.2 Gradualism, Transitivity, and Sorites

PI seems to be all or nothing such that there is a "yes" or "no" answer to whether P_1 at $t_1 = P_2$ at t_2. But the Complex View entails that persistence is gradual and partial. This raises three problems. First, a gradual persistence relation is intransitive, and this leads to an odd result. Suppose P_1 at t_1 is appropriately similar to P_2 at t_2, P_2 at t_2 is appropriately similar to P_3 at t_3. It would seem to follow that PI is sustained for P_1 and P_3. But that may not be the case since P_1 and P_3 may not be *appropriately* similar.

Second, this difficulty raises a sorites problem: For any P_1 at t_1 and P_n at t_n later than $t_{1,}$ if P_1 and P_n are exactly similar, then P_1 persists. What if P_n is exactly similar to P_1 except for the loss of one memory? Surely P_1 persists. If we iterate the question many times with each query involving the loss of one more memory, at some point, a cut-off point must be reached such that out of myriads of memories, the loss of just one more memory causes PI to cease for $P_{1.}$ The problem is, of course, that it seems metaphysically unacceptable in this context for such a miniscule cause to have such a drastic effect. And any such cut-off point seems arbitrarily drawn.

Third, a way to solve the sorites problem is to say that for any subsequent person state to P_1 at t_1, there just is no fact of the matter involved. Whether or not persistence obtains is a decision, a social construction based on our interests. This implication may fairly be taken as a *reduction* for the gradualism problem. It is an inescapable fact as to whether one literally continues at some latter time.

7.4.2.3 The Fission Problem

The persistence conditions for PI can be one to many. Familiar and widely employed thought experiments show this to be the case in many ways. For example, if a fission case presents us with two psychological duplicates who are equal candidates for the persistent person before the fission event, we have a problem, especially if we wish to avoid claiming the choice is arbitrary. A familiar strategy is to use the “no closest continuer” requirement. Stated differently, if x is the original person and y, z are the two post-fission duplicates, and we want to know if x persists as y, the “only x and y” principle applies. As is well known, the question of x’s endurance or persistence as y is one involving only x and y. z has no relevance unless y and z are internally related, which they clearly are not (y could die and z live).

---

[61] See Swinburne, R. (1997). *Evolution of the soul* (rev. ed., pp. 147–160). Oxford: Clarendon; *Mind, brain, and free will*, pp. 150–158.

Another strategy is to identify the correct persistent fission product as the one whose psychological traits are sustained by the same brain that belonged to x. But besides the fact that all the problems cited in the previous argument apply with equal force to brains, there are additional thought experiments that involve y and z each getting exactly half of x’s brain (and body), so appealing to brains just kicks the can further down the street.

Advocates of this view must tell some sort of story to solve fission cases and this will most likely include identifying the appropriate kind of relation, usually causal, between or among the relevant person stages. To accomplish this task, it is quite popular these days to expand the reductive analysis above into a thesis of the sortal dependency of PI:

> *Sortal Dependency of PI*: $(x_1 \text{ at } t_1 = x_2 \text{ at } t_2) \rightarrow [\exists S(x_1 \text{ at } t_1 =\mathrm{S}= x_2 \text{ at } t_2) \wedge (x_1 \text{ at } t_1 =\mathrm{S}= x_2 \text{ at } t_2) \leftrightarrow \mathrm{R}x_{1\times2}]$ where S is a sortal term and, *sans* bodily views, R is a psychological relation of some kind for persons (e.g., having the “same” memories as).

In general, R will be spelled out in light of the sortal concept expressed by “S,” e.g., being a statue, a dog, a driver’s license, and so on.

David Wiggins is among those philosophers who, in the strict, philosophical sense, apply sortal terms to Aristotelian substances only, and take S to express in the ideal case a real natural kind essence that generates an appropriate R for establishing genuine endurance. But when Wiggins applies this approach to aggregates, e.g., natural formations, artifacts, he takes the approach merely to be ways of taking things relative to our interests, with the result that mere persistence is established.[62] However, most follow Harold Noonan, who takes the sortal-dependency move to involve nominal essences that merely yield persistence conditions useful for various purposes.[63] Since this approach smacks of conceptualism (a problem with which Wiggins himself wrestled) and constructivism, we set it aside in search of a more promising approach for the psychological Complex View.

7.4.2.4 Immanent Causation to the Rescue?

In our view, that approach adds to the causal chain view of R another condition, viz., that Immanent Causation (IC) is at work in the chain. And the most rigorous account of IC has been offered by Dean Zimmerman.[64] Unfortunately, Zimmerman's account of IC is not particularly relevant to our present discussion, so we set it aside.[65]

---

[62] Wiggins. (2001). *Sameness and* substance. Cambridge: Cambridge University Press, chapters 3–5.

[63] Noonan, H. Personal identity and perplexities. In Gasser and Stefan (pp. 82–101).

An alternative—and more relevant—view of IC is one that was advanced by the late Medieval Aristotelians according to which when a power or faculty acts and terminates within itself (strong immanence) or within the substance to which it belongs (weak immanence), it is a case of IC. Thus, in discussing organisms, David Oderberg characterizes IC as "causation that originates within an agent and terminates in that agent …"[66] An action directed entirely outward, that terminates on or in something other than the acting substance, is a case of transient causation.[67]

Kevin Corcoran advocates this classical understanding of IC as a solution to the problem of PI.[68] According to Corcoran, a necessary condition for PI is the literal endurance of the constituting body/animal of a person. This plus some analysis of being a person, e.g., a relational one, is sufficient for PI. Corcoran claims that the living body remains numerically the same over time despite part alteration if or in the event that the different stages in the body's career stand in an IC relation to each other in the right way. Corcoran's account applies to organisms/living bodies/animals as genuine substances and IC grounds strict identity through the relevant entity's temporal stages.

Does this account work? For three reasons, we think not. First, this IC account is not necessary for PI. Consider what is surely a possible world w, an exact duplicate of the actual world, except for one feature: In w, divine occasionalism is the only causality that occurs. Specifically, substances do not exhibit either immanent or transuent causation. Still, substances, e.g., human persons, are genuine enduring continuants in w through various kinds of accidental change.

Second, the IC account is circular and presupposes as fundamental the very thing that is supposed to be derivatively grounded by IC. For a causal action to count as immanent, the action must originate and terminate within the same substance. So, IC is grounded in the more fundamental sameness of substance. Thus, IC cannot serve as the ground for the very thing it presupposes in the first place to count as IC.

---

[64] Zimmerman, D. (1998). Immanent causation. In J. E. Tomberlin (Ed.), *Philosophical perspectives Vol. 11: Mind, causation, and world* (pp. 433–471). Atascadero, CA: Ridgeview Publishing Company.

[65] Several features of Zimmerman's account are not of importance for dealing with MAs, dynamic systems, or animalist construals of organisms: 1. He seeks an account of mere persistence, not endurance. 2. His focus is on homeomerous K-masses involving locutions like "some K" or "the K" such that every mass of homeomerous stuff-kind is K through and through, i.e., every proper part of K is a mass of K. It may be that "some water," "some gold" or atomic simples are to be construed as homeomerous masses, but organisms or persons construed as aggregates are not. Moreover, virtually no Mere Substance Dualism takes the person to be composed of soul-stuff. At one time, Richard Swinburne seems to have held to such a view (see his *Evolution of the soul*, pp. 153–155), but he subsequently abandoned it; see *Mind, brain, and free will*. (2013). (pp. 32–39, 149). Oxford: Oxford University Press. 3. He grants that ME is basically true for homeomerous masses, and, thus, they suffer the same problems of diachronic persistence as MAs. 4. He cashes out his view of causation in terms of nomic subsumption of events, very similar to D. M. Armstrong's position. However, we employ a causal powers framework.

[66] Oderberg, D. S. (2018). The great unifier: Form and the unity of the organism. In W. M. R. Simpson, R. C. Koons, & N. J. The (Eds.), *Neo-aristotelian perspectives on contemporary science* (pp. 211–233). New York, NY: Routledge. The quote is from page 211.

[67] Cheyne, D. D. (2000). *Life's* form (pp. 58–59, 121). Ithaca, NY: Cornell University Press. Cf. Thomas Aquinas, *ST* IqI8a3, especially *Reply Obj. 1.*

[68] Corcoran, K. (2006). *Rethinking human nature* (pp. 72–73, 127–130). Grand Rapids, MI: Baker.

Finally, the account conflates two questions: (1) What causes P_1 at t_1 to exist at t_2? (2) What accounts for P_1 at t_1 being identical to P_2 at t_2? IC might—might—be a good answer to (1), but it does not follow that it is a good answer to (2). Indeed, it fails in this regard.[69]

7.4.3 Two Physicalist Alternatives to Human Persons as MAs

We conclude this chapter with a brief critique of two physicalist depictions of human persons that are rivals to the mereological-aggregate position and that seek to defend endurantism.

7.4.3.1 We Are Atomic Simples

7.4.3.1.1 Chisholm on Extended Physical Atoms or Simple Souls

Though not widely embraced by strict physicalists, Roderick Chisholm is well-known for having suggested that we persons might be fundamental atomic simples of some sort.[70] According to Chisholm, whatever we turn out to be, we are fundamental substances and not modes or some other sort of derivative, dependent entity. We are also enduring continuants and not *entia successive*. Given this, there were two basic views Chisholm proffered as serious contenders for our nature and identity: (1) monads—simple (without proper substantial parts, unextended/ indivisible) immaterial souls—and (2) some unique kind of physical particle or piece of matter of a microscopic nature not recognized by contemporary physics, a proper part of the brain, a spatially extended physical substance that thinks and wills. Let us call these the Simple Soul (SS) View and the Extended Physical Atom (EPA) View, respectively.

How do we decide which view is better and what one's persistence conditions are? Chisholm says we should begin by considering the nature of our mental properties with which we are directly acquainted.[71] When we do this, we will conclude that, necessarily, the only kind of thing that could exemplify them is a substance, and, possibly, only a simple substance. Moreover, mental properties have five formal characteristics that, taken together, are the ingredients for what it is to be a qualitative property. The five are:

> *Open*: Open in the sense that a property is open *iff* it is possibly such that, for any numbers m and n, there are n substances that have it and substances that do not. It is possibly such that mental properties be exemplified by simple substances. Yet they are "open" to being exemplified by any number of substances and by just that number.
>
> *Repeatable*: Repeatable in the sense that, possibly, there is something that does not have it but did or will have it.
>
> *Non-Compositive*: Not compositive where a property is compositive *iff* it is necessarily such that any compound thing that has it has a proper part that has it. Here the properties of the proper parts also characterize the whole.

---

[69] For a list of other arguments against the complex view, see Madell, G. (1981). *The identity of the self* (pp. 1–22). Edinburgh: Edinburgh University Press.

[70] See Chisholm, R. M. (1986). Self-profile. In R. J. Bogdan (Ed.), *Roderick M. Chisholm* (pp. 3–77). Dordrecht: D. Reidel Publishing Company, especially 71–77.

[71] For a presentation and evaluation of the argument to follow, see Quinn, P. L. (1997). Tiny selves: Chisholm on the simplicity of the soul. In L. E. Hahn (Ed.), *The philosophy of Roderick M. Chisholm* (pp. 55–67). Chicago and La Salle, IL: Open Court, especially 58–60.

Non-Divisive: Not divisive where a property is divisive *iff* it is necessarily such that any compound thing that has it has a proper part that has it. Here the properties of the whole (e.g., being extended) transfer to a proper part.

Internal: Internal (i.e., intrinsic) properties of substances where such is the case *iff* necessarily, it is such that whatever has it is a substance and either it is necessary to whatever has it or it is necessarily such that whatever has it has every open and repeatable property it implies. If a thing has an internal property, that does not tell us about any other thing.

7.4.3.1.2 Mental Subjects Are Simple Souls: Assessing the Case

Based on this, Chisholm argues that every qualitative property is possibly exemplified by a simple substance and no qualitative property is known to be possibly exemplified by a compound substance. According to Chisholm, purely psychological properties are not "part implying," that is, the statement that a purely psychological property is exemplified does not logically imply that the thing having that property has proper parts.[72]

First, let's consider the EPA view. There are several problems with this alternative.

Problem 1: It makes the self metaphysically divisible, and in Chapters 5 and 6, we offered reasons for doubting that this is possible.

Problem 2: It suffers from serious difficulties in proffering a self that cannot ontologically ground or explain the sort of synchronic unity of consciousness we possess (see Chapter 6).

Problem 3: Suppose S is an extended atomic particle. Suppose further with Chisholm, that for some object O to be spatially extended, O must have proper parts at different locations within the extended length. Why think S is the self instead of a proper part of S? And why think the self is not a proper part of that proper part of S? Either this generates a vicious ontological regress (since at each stage of analysis, an equally good candidate for the self as the analysandum reoccurs in the analysans and no progress is made) or one selects an arbitrary stopping point at S itself or at one of its proper parts and announces that it is the self.

Problem 4: It requires either arbitrarily singling out one physical particle among the many trillions in the brain as conscious or assigning to it a special status.

Problem 5: It postulates such a bizarre entity that most physicalists both reject and acknowledge that many of the arguments for physicalism (e.g., correlations as a ground for identity between various mental states and identifiable brain states) are not available for the EPA view.

But what if we take the spatially extended particle to be bereft of proper parts, and be some sort of mass, e.g., atomless gunk, a homeomerous K-mass, or a homogeneous continuum? We don't think this provides an undercutting defeater for *Problem 3*.[73] The reason is that as masses, these entities are subject to metaphysical division, they are degreed entities, and the iteration expressed in *Problem 3* could, with adjustments, be run against S by asking why S as opposed to 99% of S.

---

[72] Chisholm, Self-profile, p. 73.

[73] Hud Hudson discusses several different ways there could be an extended material simple (no proper parts). See his Simples and gunk. (2007) *Philosophy Compass, 2/2*, 291–302. Interestingly, he claims that "gunk" is a count noun, not a mass term. But as a count noun, we can still properly talk of a hunk or piece of gunk instead of what he calls the awful "a gunk." In this case, a piece of gunk could still be trimmed down to, say, 95% of that piece.

One could respond that besides homeomerous K-masses, if we remove some mass from S, we may run into a different sort of mass—e.g., a proper part of gunk—that is not the right sort of matter as is S to have mental properties. But this won't work. Besides contingently correlating mental and physical properties/events, it is hard to identify exactly what sort of matter—stuff or atomic simple—is needed for emergent mental properties. After all, there are other possible worlds that count physical worlds, but which have different kinds of matter and laws from the actual world, as well as emergent mental properties.

The SS position is not vulnerable to the four objections listed above. And in its favor is Chisholm's point that we know that, possibly, mental properties are exemplified by an SS. It seems that none requires extension to be exemplified. By Ockham's razor, compound or primitively extended objects are not needed.

7.4.3.1.3 The Ubiquitous Rejection of the EPA Position

Two final points. To begin with, ME applies to all the versions of the EPA position, but not to the SS view. So, we could run a Plantingian Replacement argument where the person is aware of being an enduring continuant while reading the newspaper even though mad scientists are slowly drawing mass out of the extended brain particle and simultaneously replacing it with the same sort of mass, all the while keeping the total mass constant. Moreover, at the end of the day, a choice between the two may be less significant than one might think. The self in the EPA view is a psycho-physical entity and there are versions of hylomorphic dualism that take the soul *qua* substantial form to be simple and capable of disembodiment, but also to possess the dispositions for creating, unifying, and informing a body.[74] The difference between the two involves the matter of extension and the direction of ontological grounding. But in either case, the entity identified as the self should make a staunch naturalist and strict physicalist a bit nervous.

At the end of the day, the EPA view has not been well received. Richard Taylor comments,

> This is, of course, not a conception of a person to which anyone would naturally be drawn. Indeed, Chisholm is probably the first person in the history of the world to suggest it, and he conceded that it invites ridicule.[75]

Almost no one has considered Chisholm's view, much less found it plausible.[76] The few who mention it do so with ridicule. Physicalist subject simplicity is, in the words of

---

[74] Robert Koons lists seven versions of staunch hylomorphism and adopts a different view that amounts to a combination of three of the seven. See his Staunch vs. Fainthearted hylomorphism: Towards an aristotelian account of composition. (2014 April). *Res Philosophica*, *91*(2), 151–177. One of us has offered a different view that is called a "Thomistic-like" position. See Moreland, J. P. (2018). In defense of a Thomistic-like dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 102–122). Oxford: Wiley-Blackwell. For an example of faint-hearted hylomorphism, see Koslicki, K. (2008). *The structure of objects*. Oxford: Oxford University Press.

[75] Taylor, R. (1997). Chisholm's idea of a person. In L. E. Hahn (Ed.), *The philosophy of Roderick M. Chisholm* (pp. 45–46). Chicago, IL: Open Court.

[76] See, e.g., Merricks, T. (2001). *Object and persons* (p. 111). Oxford: Oxford University Press.

Andrew Bailey, “inconsistent with (any plausible form of) materialism.”[77] Peter van Inwagen calls the view “absurd on both empirical and conceptual grounds.”[78] Ultimately, Chisholm may well have rejected the view, arguing instead that the subject of consciousness is a monad, an immaterial, indivisible, simple substance. But it is more likely that he left the matter undecided.[79]

7.4.3.2 We Are Homeodynamic Systems

Virtually everyone recognized that there is some sort of special synchronic and diachronic unity to living organisms. The question is: How are we to analyze and account for this? As Christopher J. Austin and Anna Marmodoro note:[80]

> [A]lthough composed of an uncountable number of constituents, each of these multi-layered microcosms [i.e., organisms] is a fundamentally unified being—each is in some way *one*, rather than *many*. But in virtue of what, metaphysically, are organisms more than merely bundles of biological bits whose diachronically disparate collections are continually washed away in a Heraclitan flux? In other words, what secures, metaphysically, an organism’s continued persistence *as one* over time?

It seems clear that staunch naturalists and strict physicalists operating in light of the standard mereological hierarchy lack the ontological resources to address these problems. But with adjustments, two additional versions of physicalism besides the atomic simple view have sought to solve these unity problems without employing something “spooky” like a soul or a classic substantial form (SF) employed by staunch hylomorphists: the homeodynamic systems view and animalism (see Chapter 6 for our treatment of animalism). We begin with the former by examining a representative exposition of Nancey Murphy’s view.[81]

The systems approach does not jettison the standard mereological hierarchy since it entails that at higher levels of complexity, the existence in complex wholes of novel properties, dynamic processes and laws require a holistic-systems perspective not needed at lower levels. According to Murphy, the reduction of the causality of wholes to that of their parts is false for complex systems. To see why, we must understand the type of downward causation (wholes to parts) in such systems. They reside at higher levels of the mereological hierarchy below which levels are related by the standard separable part–whole relation, objects are MAs composed of separable parts standing in external relations to other such parts, and these parts are ontologically prior to their composed wholes.

There are nonbiological and biological examples of complex systems with the latter including organisms and, more specifically, human beings (which Murphy seems to use interchangeably with whole persons). Focusing on human beings, Murphy asserts that there is a progressing convergence between physicalism and cognitive neuroscience in this way: All human capacities once attributed to the soul are now "fruitfully" studied as brain processes; mental processes are now being "related to" neuron functions/systems. Not only is the brain "involved in" specific mental operations, but very specific mental operations are being "correlated with" highly specific brain regions. These findings place a burden of proof on SDs who want to postulate an additional entity (a soul) that is unneeded and that fails to satisfy epistemic simplicity.

---

[77] See, e.g., Bailey, A. (2014). You needn’t be simple. *Philosophical Papers*, *43*(2), 147.

[78] van Inwagen, P. (1990). *Material Beings* (p. 189). Ithaca, NY: Cornell University Press.

[79] Chisholm, Self-profile, pp. 73, 77.

[80] Austin, C. J., & Marmodoro, A. Structural powers and the homeodynamic unity of organisms. In *Neo-Aristotelian perspectives on contemporary science* (pp. 169–183). The quote is from page 169.

[81] Murphy, N. C. (2018). For nonreductive physicalism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (pp. 317–327). Oxford: Wiley-Blackwell.

Murphy endorses van Gulick's claim that complex systems are patterns that maintain themselves over time even with part replacement, and she points to whirlpools as an example. Higher-level patterns have a degree of independence from their underlying realizers and can exert top-down causation without altering the laws of physics. Against this claim, causal reductionists raise the following dilemma: Either this causation involves some unusual means or some "spooky" entity (e.g., a substantial form?). If the former, then these interactions must be governed by ordinary physical laws and all causality is bottom-up. If the latter, then this is scientifically unacceptable. Moreover, continues the causal reductionist, even if we grant top-down causation, there is still no agency for humans since we are still determined by our parts and top-level environmental factors.

For Murphy, the answer to the objections is found in systems theory and the emergence of agency. Her position consists in five fairly representative theses relevant to our discussion.

(i) Give up the bias for things and their intrinsic properties for processes and relations. For example, a mammal is composed of a circulatory and reproductive system, not of carbon and hydrogen.

(ii) Systems have permeable boundaries with a tighter coupling of boundary components compared to inner coupling that allow transport of matter, energy, and information.

(iii) Systems are not MAs. Regarding systems, the properties of their components—their nature and identity—depend on their parts being in the system and the relations among constituent processes (and parts?) are internal with respect to each other.

(iv) Systems range from great diachronic stability to wild fluctuations because such complex systems are nonlinear. Their degree of stability is determined by degree of stability to variations of initial conditions and to the extent that feedback processes dampen fluctuation. Systems at the edge of chaos are adaptive in that they have the freedom to explore new possibilities, may jump to a new, higher organization, and are characterized by goal-directedness/irreducible teleology along with self-maintenance.

(v) Complex adaptive systems imply a new concept of causality. Efficient causality is inadequate. Rather, we need to think of the relations among constituents of systems as constraints that reduce the number of possibilities that can happen. And since constituents of systems stand in internal relations to each other, a change in one brings about a change in another. Constraints are not things that push other things; rather, they are internal relations among constituents of the system.

7.4.3.2.1 Objections to the Systems Approach

Does Murphy's employment of what it is to be a complex system succeed in providing a metaphysical resource for grounding agency and synchronic/diachronic unity that we all find it almost impossible not to attribute to ourselves? We do not believe it does, and in our response to follow, we sometimes limit our discussion to human persons and offer six responses to Murphy's systems theory.

7.4.3.2.1.1 What We Seem to Know About Ourselves

For one thing, one finds oneself to be a *thing* in a properly basic way, not a system, especially not a system that is built up out of processes. As E. J. Lowe notes, along with common sense and ordinary language, "I feel *myself* to be some *thing,"* in particular, a simple thing."[82] Along similar lines, Frank Jackson acknowledges: "I take it that our folk conception of personal identity is Cartesian in character—in particular, we regard the question of whether I will be tortured tomorrow as separable from the question of whether someone with *any* amount of continuity—psychological, bodily, neurophysiological, and so on and so forth—with me today will be tortured."[83]

Moreover, it is evident to me *that I am not an ontologically vague object* like a cloud (since we do not believe in vague objects, we take clouds to be water vapor arranged cloud-wise). *I have distinct, continuous boundaries*. But complex systems have permeable boundaries that are porous and gappy. Further, as we have argued, if we are enduring continuants, then we are wholly present at each time we exist; we move wholly through time and *do not have extended temporal parts*. But complex systems construed as processes or composed of processes have none of these features. Specifically, processes are temporally extended. Thus, we are not complex systems.

7.4.3.2.1.2 Avoid Spooky Entities at All Costs

Ubiquitous throughout defenses of physicalism is the exhortation to proffer an account of human persons that avoids quantifying over "spooky" entities. The term packs rhetorical force by its association of SDs with those who believe in ghosts and other superstitious, scientifically discredited entities. But the exhortation is more than rhetorical. It also expresses the importance of staunch naturalism and at least weak scientism in placing constraints on positions and methods adopted in philosophy of mind. Thus, in proposing views in philosophy of mind, one should avoid postulating souls, substantial forms, top-down causation that intervenes in the laws of physics/chemistry, animal spirits, etc. Murphy is explicitly committed to this exhortation. As an aside, this exhortation follows hard on her rejection of first-philosophy and, along with it, any distinctive philosophical contribution a philosophy of human nature could supply.

So, it is odd to say the least that her account includes internal relations and irreducible teleology. As far as we know, all the physical (non-spooky) relations in the hard sciences and that constitute wholes and structural properties that populate the mereological hierarchy are external. Thus, internal relations seem to be spooky indeed and they are instantiated "all at one go," states of affairs that are utterly inexplicable. And irreducible teleology is surely a discarded notion on the trash heaps of scientific advance. Please note that we accept irreducible teleology. But we do not think that the meager ontological resources of Murphy's non-reductive physicalism provide a rich enough framework for Murphy to help herself to irreducible teleology (or internal relations).

One could respond that by teleology Murphy is merely describing sequential process in a teleological way, but since her appeal to goal-directed activity does ontological work in sustaining and directing diachronic development of systems through probability space, this response seems wrong. The problem here is that teleology itself, especially construed as a metaphysical principle of activity, has long been regarded as having gone the way of the dodo and phlogiston. Such a principle becomes even more spooky if, as is usually the case, it is understood as the diachronic aspect of a substantial form.

---

[82] Lowe, E. J. The probable simplicity of personal identity. In Gasser & Stefan (pp. 137–155). The quote is from page 152.

[83] Jackson, F. (1998). *From metaphysics to ethics* (p. 45). Oxford: Clarendon.

It is unclear how Murphy decides which spooky entities are allowed and which are not. At the very least, we are owed some sort of account here and, to our knowledge, she has never produced one.

7.4.3.2.1.3 Complex Systems as Synchronically Unified Wholes

As a paradigm advocate of systems theory, Murphy must account for the unique unity of organisms—especially human beings—*qua* biologically complex systems from within a rather meager ontology that attempts to stay as close as possible to the ontology of the hard sciences. On the one hand, this requires her to avoid substantial forms and their role as synchronic unifiers and as grounding diachronic principles of teleological activity. She explicitly claims that, given advances in neuroscience, there is a significant burden of proof on a substance dualist and this burden would apply to substantial forms. On the other hand, she must avoid reducing complex systems to MAs with an implicit commitment to a bottom-up, part-priority mode of explanation for the features and activity of those systems.

Given these constraints, we believe there are two undercutting defeaters for her account of the nature and ground of the system *qua* unified whole. Recall that her systems approach requires replacing parts/things with intrinsic properties in favor of composing processes, relations, and subsystems. Further, a new concept of causality must replace efficient causes with the imposition of causal constraints involving regulatory, teleological processes through possibility space. These processes achieve diachronic self-maintenance and unity.

We have already examined teleology above and will not repeat our evaluation here except to make one new point. Murphy speaks of these processes in the plural and that is the correct way to describe them. An organism exhibits numerous processes and systems at and through time. What is left out of this account is an explanation for why these many processes/systems are subsumed under and co-opted to serve a unified one, viz., the organism.

In response, one could appeal to Murphy's claim that "[s]ystems theory takes the relations among the constituent *processes* of a system to be *internal*."[84] It seems that these internal relations standing between various processes as relata are what ground the synchronic and diachronic unity of the whole. But this response does not solve the unity problem; it merely postpones a solution. Now, the question arises as to what unifies all these internal relation instances into one unified system. Recall that an internal relation relates only its relata. That's all. But as Murphy's statement reveals, a homeodynamic system has numerous constituent processes. Whether we allow processes or parts to be fundamental, either way we have a large plurality of entities. Internal relation-instances will also be large pluralities that stand between or among different relata. One is still left with the problem with which we began: Why or how is this plurality of distinct internal relation-instances brought together into a single unified system? *Sans* substantial form what could possibly ground this phenomenon? To our knowledge, there is no good answer to these questions.

---

[84] Murphy, For nonreductive physicalism, p. 323.

We see another serious problem with the systems-theory account of the organism *qua* system as a new synchronically unified whole. Granting that systems are not aggregates in that the former are new wholes “over and above” their micro- and macro-parts and the various processes/systems that compose them, then if we ask from where that new whole came, we seem to get some sort of creation *ex nihilo*. Part of the problem here is that property emergence cannot be made coherent if the subvenient base is taken to be a sufficient condition for such emergence, precisely since this approach requires base-level, brute physical entities to generate something of an entirely new kind “all at one go” of which the base entities are bereft. This is especially true when the emergent properties are irreducibly, intrinsically mental. This is creation *ex nihilo*.

The same difficulty counts against the emergence of a new composite whole such as one of Murphey’s organismal systems. As Antonella Corradini observes, a systems view applied to emergent individuals forces us to contemplate “a composite physical system’s giving rise, all in one go, to a whole, self-contained, organized system of properties bound up with a distinct individual.”[85] In addition, Murphy’s view also entails that at the right level of complexity, external relations are replaced with internal ones, the nature and identity of separable parts are replaced with inseparable parts with a new nature and identity, and efficient causality is replaced with causal constraints.

Systems theorists like Murphy seem to be faced with two unacceptable horns of a metaphysical dilemma. On the one hand, they can bite the bullet and assert that at some point a purely physical system gives rise to a novel kind of whole. In this case, creation *ex nihilo* seems to be a defeating entailment.

On the other hand, they can claim that a system’s constituents (e.g., separable parts, properties, events/processes) undergo an essential change into inseparable entities that stand in internal relations to each other. Recall that relata can stand in internal relations only if they have the proper kindedness that grounds this. Thus, the relata must change their nature ontologically prior to shifting from standing in external to internal relations. In this case, *sans* a substantial form, the problem is to explain how such a thing is metaphysically possible, what the underlying subject of change is, and why the new whole exhibits a *per se* unity that goes beyond the accidental unity of an MA and is usually associated with a classic substance. Again, we are unaware of adequate answers to these questions.

These utterly mysterious metaphysical puzzles can be easily avoided. Focusing on a mental subject, the emergentist must hold that the base-level entities are necessary but not sufficient, though we will see reasons in Chapter 9 to doubt this. For a mental (property or) substance to emerge from a biological base, there must be a nonmaterial entity with ontological independence (to exhibit real downward efficient causation) that exists from the very beginning of the emergent process. The clear candidates for such an entity are a substance (a soul) or a substance-like entity (a substantial form) to unify the organism at and through time, to account for the shifts that take place at a certain level of complexity, and that avoid creation *ex nihilo* of a new compound entity/system by having the mental substance there from the beginning. If we take emergent properties to be *sui generis* in nature, to exhibit new causal powers, and to arise from and be sustained by the underlying base, then this account eschews emergent properties in favor of a new substance that actualizes new properties it exemplifies at different stages of development. And a single substantial soul or substantial form grounds the unity of all of an organism's parts (by transforming them into inseparable parts or modes), systems, and processes. More on this shortly.

---

[85] Corradini, A. (2008). Emergent dualism. In A. Antonietti, A. Corradini, & I. J. Lowe (Eds.), *Psychophysical dualism today: An interdisciplinary approach* (pp. 185–210). Lanham, MD: Lexington Books. The statement is on p. 206. Corradini is citing O’Connor’s view to use it as a foil for her own version of substance dualism.

7.4.3.2.1.4 Problems with Harmonizing Physicalism and Downward Causation (DC)[86]

There is a great deal of diversity among systems theorists regarding (i) the nature of the cause in DC (principles/laws, wholes, boundary conditions, context-sensitive constraints, patterns of organization, processes, entities); (ii) the object of causal influence (parts/constituents of a whole or system, lower-level properties, events/actions/processes, conditions); and (iii) the nature of DC itself (e.g., laws/regularities, efficient causality, the placing of lower-level causal constraints). This diversity may simply be a result of different thinkers developing different depictions of what is going on, and there is nothing more to it. However, the variety may reflect a deeper problem that has not been adequately addressed. Regardless of the correct way to understand why such diversity exists, there is still a deep problem with physicalist versions of systems theory.

Put simply, there is a substantial and widely acknowledged difficulty in attempting to harmonize physicalism, along with the causal closure of the physical and the causal inheritance principle, with a view of emergent entities as *sui generis* nonphysical items with top-down causal "influence" and the unique ontological character of DC.

Marius Tabaczek correctly notes that from a scientific perspective, if philosophical reflection on physical causation must restrict causality to efficient causality due to (i) the physical principle of action and reaction (in principle) amenable to mathematical description; (ii) this principle has dominated natural science since Modernity.[87] In light of this, the fundamental difficulty in harmonizing genuine emergence and DC construed as efficient causation and the causal closure of the physical has been nicely expressed by van Gulick:

> The challenge of those who wish to combine physicalism with a robustly causal version of emergence is to find a way in which higher-order properties can be causally significant without violating the causal laws that operate at lower physical levels. On one hand, if they override the micro-physical laws, they threaten physicalism. On the other hand, if the higher-level laws are merely convenient ways of summarizing complex micro-patterns that arise in special contexts, then whatever practical cognitive value such laws may have, they seem to leave the higher-order properties without any real causal work to do.[88]

---

[86] Our presentation of this objection is indebted to Tabaczek, M. (2019). *Emergence* (pp. 71–78). Notre Dame, IN: University of Notre Dame Press.

[87] Tabaczek, p. 76.

[88] van Gulick, R. (2007). Who's in charge here? And who's doing all the work? In N. Murphy & W. R. Stoeger (Eds.), *Evolution and emergence: Systems, organisms, persons* (pp. 64–65). Oxford: Oxford University Press.

The problem with responding to this problem by appealing to selective activation of lower-level causal powers, provision of causal constraints, and so forth, is this: Insofar as such postulations involve exerting causal influence, the dilemma does not go away. Causal influence seems to be very close to efficient causality, but even if it is genuinely distinct, causal influence must make a difference in the direction of physical events. If such influence is physical, then reduction is hard to avoid. If it is not physical, then closure is violated by something "spooky." Moreover, since our focus in this book is human persons, free actions understood in compatibilist or libertarian ways involve efficient causality at some point. In Chapter 10, we will see that this is clearly the case with the libertarian notion of an exercise of active causal power.

7.4.3.2.1.5 Complex Systems as Enduring Continuants

Finally, we believe systems theory does not have the ontological resources to ground the diachronic sameness for complex systems that we human persons enjoy. What could possibly serve as such a ground? Processes? No, because processes exist at different times by having different temporal process-slices (or intervals) at different times. Entire processes do not exist through time by being wholly present at each moment at which they exist. Also, processes are constantly changing, coming, and going. As a human person develops, its various metabolic processes change, are replaced in a lawlike fashion, and develop individually (the circulation system develops new pathways as an organism grows). Further, if one has an organ removed, or partially removed as in certain brain surgeries, or loses a limb, the processes do not remain the same. They also change as one goes from being awake to being asleep. As a synchronic grouping and diachronic series of multitudes of activities constituting an organism's processes, these activities are a many in flux and, thus, inadequate to ground diachronic identity.

Systems? No, because systems exhibit the same problematic features as do processes. Structure? No, because an organism's relational structure changes over time. Further, what is needed to ground diachronic identity is sameness of structure-instance, not mere sameness of structure-type. The latter cannot individuate, but the former can. Unfortunately, structure instances are aggregates of relation-instances for systems theory, and even if a structure loses two parts related to one another and they are replaced with two exactly similar parts in the same *kind* of relation, we still have a new relation-instance, so structure becomes a series of structure-tokens as parts and other relata come and go. This point holds for external and internal relations and their instances.

Finally, systemic diachronic identity cannot be a brute fact. Brute facts are not all created equally. Thus, it is easy to see how a simple (unextended without proper parts), e.g., a bare particular, exhibits diachronic identity as a matter of brute fact. But systems are composed of multitudes of different kinds of manies—parts, processes, events, subsystems—that are constantly in flux. It is hard to see how it could be a brute fact that such systems exhibit diachronic identity.

We believe it is time to turn away from systems theory to a range of views we take to be better. As Mariusz Tabaczek bemoans,

> The central dogma of emergentism described in the previous chapter in terms of nonadditivity of causes, novelty, nondeducibility, nonpredictability, irreducibility of complex dynamical systems, entities, properties; ontology of emergent levels of complexity and laws governing them; and DC [downward causation] is commonly accepted by scientists and philosophers of science. But what appears to them as

> a consistent, well-thought-out, and coherent theory, may seems in the eyes of an Aristotelian metaphysician to be nothing but a confusing welter of ideas, characterized by a maddening ambiguity of terms.[89]

Specifically, we hold that extant versions of staunch hylomorphism are superior to systems theories of organisms, especially of human persons.

7.4.3.2.1.6 Versions of Staunch Hylomorphism Are Superior to Systems Theories of Organisms, Specifically, Human Persons

In an important, cutting-edge article, Robert Koons describes seven versions of staunch hylomorphism with the seventh being the one he prefers.[90] Our basic purpose is to show that several versions of staunch hylomorphism are part of the contemporary discussion. Secondarily, we will offer a difficulty for all seven versions, followed by an eighth version which is our favorite option. But if we are wrong, then our primary purpose is untouched since there are several robust positions that provide legitimate alternatives to systems theory.

Before we lay out Koons's seven versions, we must first get clear on the distinction between faint-hearted and staunch hylomorphism. According to Koons, faint-hearted hylomorphism is a version of physicalism according to which substantial form is moved from the primitive category of natural kind or the category of (essential) property to the category of relation. Thus, a substantial form is identical to a set of relations, more precisely, a structural arrangement of elements/parts of the body. The acquisition of "substantial form" is the occurrence of certain relations and the autonomous cooperation of an organism's separable parts. The resultant whole is an accidental unity and not a genuine substance. Koon notes that faint-hearted hylomorphists (e.g., Kathrin Koslicki[91]) fail to distinguish genuine substances from non-substantial wholes. The matter of an organismic whole consists in a plurality of separable parts that stand together in the relevant external relations.

Staunch hylomorphism stands in the historical stream of classic hylomorphism. There is a real difference between substances that have substantial forms and non-substances that do not. Further, substantial form is a natural-kind or species essence (or that from which the essence "flows") in the category of property or, for some, in the primitive category natural-kind. As an expression of a power's ontology of causation, a substantial form imposes fundamental powers on the substance's parts, and the nature, identity, and causal powers of those parts at least partially come from the substantial form of the whole. Moreover, the substantial form grounds the synchronic unity of the substance it informs, as well as the teleological, diachronic sameness of its informed substance as it grows teleologically toward its proper end. More could be said, but this is enough for our purposes.

---

[89] Tabaczek, p. 63.

[90] See Robert Koons, Staunch vs. fainthearted hylomorphism: Towards an Aristotelian account of composition. For a concise and informative treatment of Aquinas's view of the human being, see Feser, E. (2018). Aquinas on the human soul. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 88–101). Oxford: Wiley-Blackwell.

[91] See K. Koslicki, *The structure of objects.*

Koons lists seven versions of staunch hylomorphism with his preference being (7):

1. *Aristotelian Parts-Nihilism*: Emergent substances have no actual parts at all.
2. *Reverse ME*: Parts of substances cannot exist except as parts of substances of the same kind.
3. *Downward sustenance*: The persistence and operation of the whole substance *cause* the persistence of its parts.
4. *Upward sustenance*: The persistence and cooperation of the substance's parts *cause* the persistence of the whole.
5. *Upward power migration*: Some (or all) causal powers *migrate* from parts to the whole.
6. *Teleological subordination*: The powers and activities of the parts are teleologically ordered to some end pertaining to the whole.
7. *Parts as sustaining instruments*: The constraints of options 4, 5, and 6 are all jointly actualized.

Our purpose in presenting these is to show that several versions of staunch hylomorphism exist, and to claim that precisely because they express a much richer, detailed and pervasive ontology than do systems theories, they have better resources for grounding synchronic unity, diachronic sameness, and teleological development toward a proper end set by an organism's essence (or substantial form). Still, it may be worth stating our own preferences about staunch hylomorphism without denying that all are legitimate rivals to systems theory.

To begin with, we agree with Mariusz Tabaczek when he says that

> … Koons fails to recognize the Aristotelian notion of primary matter, with the result that he defines the material aspect of hylomorphism in terms of mereologically understood elementary particles. Consequently, in light of classical Aristotelianism, his argument against options 1 and 2 (that Aristotle's theory assumes the reality of material (mereological) substrates that endure through any substantial change) does not stand. For … what persists through each and every substantial change is primary matter.[92]

---

[92] Tabaczek, p. 223. Prime matter is far from an outdated idea and finds a contemporary analogue in bare particulars which Gustav Bergmann called "Aristotelian prime matter splintered." See Bergmann, G. (1967). *Realism: A critique of Brentano and Meinong* (p. 110). Madison, WI: University of Wisconsin Press; cf. 25, 77. For a defense of bare particulars, see Moreland, J. P. (1998). Theories of individuation: A reconsideration of bare particulars. *Pacific Philosophical Quarterly*, *79*, 251–263. Jeffrey Brower has proposed a gunky, massy stuff view of prime matter. See Brower, J. (2015). Matter. In *The Cambridge dictionary of philosophy*. New York, NY: Cambridge University Press; *Aquinas's ontology of the material world: Change, hylomorphism, and material objects*. (2017). Oxford: Oxford University Press. According to Brower, stuff is not an individual thing, but a mass that can be measured and exist in greater or lesser degrees of quantity. Brower's position seems to entail the existence of atomless gunk, which we reject. For a brief critique of atomless gunk, see Koons, R. C., & Pickavance, T. (2017). *The Atlas of reality* (pp. 494–497). Malden, MA: Wiley-Blackwell. Further, it would seem that a change in individuator is a sufficient condition for a change in the individuated. And since human persons are nondegreed, all or nothing kinds of entities, and masses/stuffs can exist in different quantities and, thus, "be identical to" an original mass to such and such a degree upon losing some mass, it seems that massy gunk cannot individuate human persons.

Second, versions (3)–(7) violate the classic necessary condition for something being a substance in distinction from an aggregate, a bundle, or a heap. This condition has been affirmed recently by E. J. Lowe and David Oderberg and it is the claim that no genuine substance has other substances (we prefer “separable parts”) as proper parts. Not only does this principle preserve the deep unity of substance according to which its parts are inseparable modes whose identity and nature derive from the substance of which they are modes, it also affirms that the fundamental substrate of substantial change is not a set of separable parts, but prime matter.

Finally, specifying the relation of form to matter in terms of a relational organization into which the separable parts are properly placed departs significantly from classic hylomorphism in which the relation of form to matter is analyzed in causal terms with respect to the distinction between potentiality and actuality.

Given these concerns, we offer our own version of hylomorphism, what we will call *Thin Particular Hylomorphism*, that has the following advantages (some of which are shared with a few versions of (1)–(7)): (i) It solves the three problems just mentioned. (ii) It preserves the mereological simplicity of an organism, particularly, human persons, while retaining their metaphysical complexity by employing inseparable parts. (iii) It provides a framework within which the Aristotelian/Thomistic notion of essence or substantial form finds a very tight analogue with the contemporary notion of biological information. (iv) It explains (in particular, physics or chemical) part incorporation into a substance in a way consistent with the fruitfulness of employing chemistry and physics in the study of biological substantial wholes.

The human soul (hereafter, simply soul) is a simple (containing no separable parts), spatially unextended substance containing the capacities for consciousness and for animating, enlivening, and developing teleologically its body. The faculties of the soul (e.g., the mind, will) are inseparable parts of the soul containing a group of naturally resembling powers/capacities.[93] The essence (substantial form) of the soul is in the category of property and not relation—especially external relation—it grounds membership in a living being’s natural kind, it grounds diachronic identity as a principle of activity that directs an organism’s development toward its proper end, and it subsumes lower-level (e.g., chemical/physical) parts, thereby transforming separable into inseparable parts whose identity/nature derives from the substantial form of the whole. In this case, these lower-level parts undergo substantial change, but still retain (i) various qualitative, quantitative, and causal features relevant to the new whole and (ii) its original essence as a virtual, potential entity that is restored if that part is removed intact from the substance.

The late Medieval Aristotelians drew a distinction between a thick particular (the entire concrete organism; the thin particular plus accidents, including the body) and the thin particular (the essence/substantial form, the nexus of exemplification, and an individuator, in their case, prime matter).[94]

As we turn to an analysis of the body, some words by Francisco Suarez (1548–1617) provide a fitting transition:

> A pure principle of thought, however, is not suited to inform the body … Therefore, for a rational soul to be a true form of the body, it must be the principle not only of thinking but also of the operations that are exercised by the body.[95]

---

[93] Cf. Perler, D. (2015). *The faculties: A history* (pp. 3–17). Oxford: Oxford University Press.

[94] Robert Pasnau, pp. 99–114.

[95] Suarez, F. (1866). *Disputationes metaphysicae*. 15.10.25. Paris: Vives; repr. Hildesheim: Olms, 1965; originally published in 1597.

In what follows, we offer two versions of Aristotelian-style dualism that provide two different understandings of the body and the body/soul relationship: first, a strictly metaphysical thesis we shall call Metaphysical Aristotelianism (MAR); second, what we call Organicism, a more metaphysical/scientific thesis than MAR that, among other things, implies certain scientific theses that while currently in disfavor are making a comeback in recent years.

Our delineation of these two distinct Aristotelian-style views has been noted by what is most likely the most authoritative treatment of the Aristotelian metaphysics of substance in the late Middle Ages—Robert Pasnau's *Metaphysical Themes: 1274–1671*. Says Pasnau:

> [S]cholastic authors do offer metaphysical entities as principles of explanation on a concretely physical level, as efficient causes in competition with a corpuscular-mechanistic account of the natural world. The hylomorphic theory admits of an alternative formulation, however, as an explanatory schema at a different level of analysis, not competing with a corpuscular-mechanistic theory, but accounting for abstract, structural features of the world—in particular, the unity and endurance of substances … One diagnosis of the decline of scholastic thought … is that the scholastics lost their grip on hylomorphism as a metaphysical theory, conceiving of it instead as a concrete, physical hypothesis …[96]

According to MAR, a living organism is not an MA or/system composed of separable parts, bundles of properties, nor is it a concrete organism construed as some sort of whole. Rather, it is a thin particular, viz., an essence/substantial form exemplified by an individuator (usually prime matter), standing under (sub-stands) the accidental features of the organism, including its body.[97] The thin particular is identical to the organism's soul, it is mereologically simple (not composed of separable parts), metaphysically complex (containing a complex essence of properties and dispositional causal powers, exemplification, and an individuator), and it is holenmerically present throughout the organism's body (i.e., fully present to the body as a whole and fully present at each part of the body).

The thin particular played four central metaphysical roles: (1) It metaphysically grounded the special sort of synchronic unity of living things, especially compared to MAs and systems. (2) It metaphysically grounded a living thing's ability to be a continuant, teleologically sustaining strict, absolute identity through certain lawlike changes (including constituent replacement in the organism's body). (3) It provided the metaphysical ground both for placing the organism in its natural kind and for unifying that kind. (4) Its essence included an unextended structure—a blueprint and order of assembly—for building a body—an extended, informed, ordered arrangement of inseparable, functional parts, systems, and so forth that was taken to be isomorphic with that internal unextended essence. The organism's body in an essentially informed material structure modally distinct from the thin particular (the thin particular can exist without the body but not conversely since it is a dependent entity). *Sans* the soul and its substantial form or essence, the body is a corpse—an MA—and no longer a body.

---

[96] Pasnau, pp. 100–101; cf. 558–565.

[97] Ibid., pp. 99–134.

Relevant to our interaction with Murphy above, MAR advocates clearly distinguished attempts to provide an ontological classification of the nature of various capacities and their possessors from proffering an explanation of the bodily conditions required for the exercise of those capacities; and MAR advocates were clearly interested in the former, not the latter. As Dennis Des Chene points out:

> The Aristotelians, while acknowledging, even insisting, on the necessity of a material basis for the instantiation and exercise of vital powers, did not seek to reduce them to complexes of powers found also in inanimate things … For them, the project was not to find a chemical basis for life, but to describe and classify vital powers, and then, in keeping with the scheme of Aristotelian natural philosophy, to define the genera and species of living things in terms of those powers.[98]

For MAR, then, the body is key for both the functioning of the thin particular's (soul's) powers and the actualization of its various capacities. Speaking of the human soul, Des Chene observes, "The human soul is not merely joined with the body in fact. It is the *kind* of soul which, though capable of separate existence … nevertheless by its nature presupposes union with a body, and moreover with a particular kind of body, a body with organs, in order to exercise all its powers—even reason …"[99] Elsewhere, Des Chene notes: "Even the intellect requires, so long as the soul is joined with a body, a certain disposition of the brain."[100]

Thus, MAR remains consistent with, and may even entail the search for specific neurological causal/functional/dependency conditions associated with the actualization of the soul's capacities for consciousness. Such a search would not provide information about the intrinsic nature of the capacity or the property it actualizes (e.g., pain) nor about the possessor of that capacity (the soul, not the brain). But it would provide information about the bodily conditions required for its actualization. This form of dualism (as with Organicism; Section 7.4.3.2.1.6) is quite at home with the existence of contemporary neurological findings.

Moreover, while physicalism may be fundamental to a neuroscientific research program, in the specific sense in which this is true (there will be neurophysiological conditions in deep causal/functional dependency which ground the various capacities for life and consciousness), the thin particular is also fundamental to an MAR research program.

A second view among the late-Medieval Aristotelians, distinct from MAR, which we call "Organicism," has certain things in common with vitalism, though classifying it as such is a serious mistake. Pasnau notes that on this view, the soul *qua* substantial form "plays a straightforwardly causal role, explaining both the behavior and the physical structure of an animal's body."[101] In this sense, the soul is not only the formal/essential cause of the body, but also becomes (1) an internal efficient first-moving (immanent) cause of the development and structure of the body, and (2) the teleological guide for that development and structure (thus, form determines function). As with MAR, Organicism identifies the organism with the thin particular, viz., the soul.

---

[98] Des Chene, p. 7.

[99] Ibid., p. 71.

[100] Ibid., p. 96.

[101] Pasnau, p. 558; cf. 549, 552–565. There is a debate about whether or not a substantial form exhibits efficient causality. For a defense of the rejection of efficient causality, see Owen, M. (2021). *Measuring the immeasurable mind: Where contemporary neuroscience meets the aristotelian tradition* (pp. 153–156). Lanham, MD: Lexington Books. For a defense of the affirmative position, see Lecon, M. (2013). Francis Suarez on the efficiency of substantial forms. *The Review of Metaphysics, 67*(September), 107–124. See also, Oderberg, D. S. (2011). Essence and properties. *Erkenntnis*, *75*, 85–111.

Here, the soul is a substance with a spatially unextended essence or inner nature containing, as a primitive unity, a complicated, structural arrangement of capacities/dispositions for developing a body. Taken collectively this entire ordered structure is unextended, holenmerically present throughout the body, and constitutes the soul's principle of activity that teleologically governs the precise, ordered sequence of changes the substance will (normally) undergo to grow and develop.

The various subsumed physical/chemical inseparable parts and processes (including DNA) are tools—instrumental causes—employed by higher-order biological activities to sustain the various functions grounded in the soul. Thus, the soul is the first efficient cause of the body's development as well as the final cause of its functions and structure, which, along with the body's inseparable parts, are internally related to the soul's essence.[102] The functional demands of the soul's essence determine the character of the tools, but they, in turn, constrain and direct the various chemical processes taking place in the body as a whole. Organicism, then, implies the organism as a whole (the soul) is ontologically prior to its bodily parts. Along with the soul's holenmeric presence in and to the body, this understanding of the soul's essence makes it very similar to the notion of information, as used in biology today.

Many scientists take information to be a nonphysical, *sui* generis entity. While scientists can usually tell you what (nonpropositional) information does or how it is measured, they find it difficult to define. At a minimum, (Shannon) information is the reduction of possibilities.[103] If Roger tells you that he lives in California, for example, that small bit of information leaves many possibilities open. But if he gives you his town and street address, this new information eliminates a significant number of possibilities.

In addition, biologist Jonathan Wells claims that information is a *sui generis*, irreducible entity that is an immaterial, unextended, multifaceted blueprint for organismic development.[104] As such, information is present/available to the organism as a whole functioning much like a formal and final cause, and fully present/available to each cell integrated into the organism (organisms have numerous bacterial cells that, while they play an important symbiotic role with organisms, they are not, strictly speaking, "parts" of the organism). In this way, the information-in-cell C "tells" C where it is and what role it is to play with respect to the whole organism. If information is not identical to an Aristotelian essence, it seems at least to be very close.

A minority of biologists are returning to a type of Aristotelian essentialism.[105] And a growing minority of biologists are re-introducing irreducible teleology into the field.[106]

---

[102] See Kaiser, T. (2014). Is DNA the soul? Paper presented at the West Coast Meeting of the Society for Aristotelian and Thomistic Studies, June 14, 2014. Paper posted at www.aristotle-aquinas.org.

[103] Dembski, W. (2014). *Being as communion* (pp. 17–19). Burlington, VT: Ashgate.

[104] Personal email communication, Jonathan Wells (April 3, 2015).

[105] Webster, G., & Goodwin, B. (1996). *Form and transformation: Generative and relational principles in biology* (pp. 3–100). Cambridge, UK: Cambridge University Press, especially 17–25.

[106] See Toepfer, G. (2012). Teleology and its constitutive role for biology as a science of organized systems in nature. *Studies in History and Philosophy of Biological and Biomedic*, *43*, 113–119; Aranda-Anzaldo, A. (2011). *Assuming in biology the reality of real virtuality (A come back for entelechy?). Ludus Vitalis*, *XIX*, 333–342.

Moreover, an organism's parts are inseparable parts that stand in internal relations to the soul's individuated essence; they are literally functional entities constituted by their role in the whole organism. The body is developed and grows teleologically, by means of law-like developmental events, rooted in the internal essence of the soul. The first-efficient cause of the characteristics of an organism's body is its soul (containing a blueprint or information in its individuated essence); the various body parts, including DNA and genes, are important instrumental causes the soul uses to produce the traits that arise. This sort of view, along with the holism with which it is associated is also gaining ascendency in biology.[107]

To summarize, Thomistic-like organicism entails (1) the organism as a whole (the soul) is ontologically prior to its inseparable parts; (2) the parts of the organism's body stand in internal relations to the soul's essence; they are literally functional (teleologically understood) entities (the heart functions literally to pump blood); (3) the body's operational functions are rooted in the soul's internal structure (substantial form); the internal structure or essence is the blueprint, the information responsible for the body's structure and functions; (4) the body is developed and grows teleologically as a series of developmental events occurring in a law-like way, rooted in the internal essence of the human soul; (5) the first efficient cause of the characteristics of the human body is the soul; various body parts, including DNA and genes, are important instrumental causes the soul uses to produce the traits that arise; (6) the body is a mode of the soul (the soul could exist without the body but not conversely; a body without a soul is a corpse); as such it is an ensouled physical structure; thus, there are two aspects to the body—a soulish, immaterial and a physical aspect. (7) When becoming constituents of a living organism, chemical and physics parts undergo substantial change such that (i) many of their relevant qualitative, quantitative, and causal features are retained while being subsumed under and receiving their nature and identity from the organism's substantial form; and (ii) retain their original substantial form virtually/potentially such that upon separation from the organism, that form becomes actualized.[108]

We close with two final reflections. First, on this view, the role of DNA and genes can be likened to the materials needed to construct a house. Three things are needed for such a construction: specified building materials, a complete floor plan or blueprint, and a specified order in which things will be assembled. In terms of this analogy, DNA's role is to specify the patterns for making the materials (proteins) to be used in assembling the organism. Genes play a role in stabilizing certain aspects of the spatial and temporal order of growth and development, but they do not generate that order. Genes produce cell materials but not the overall plan or the internally related organization among the organism's parts. According to Organicism, living organisms are wholes irreducible to and ontologically prior to their parts.

Considerable evidence supports an Organocentric view. First, the two main functions of DNA (being copied in the process of cell division and serving as a template for protein synthesis) require the coordinated activity of numerous complex molecules and can occur only within the context of an entire cell. In fact, as Richard Lewontin says, “DNA is a dead molecule, among the most nonreactive, chemically inert molecules in the living world. That is why it can be recovered from ancient plants and long-dead animal. It has no power to reproduce itself and, while it is promoted as producing proteins, in fact proteins (enzymes) produce DNA.”[109] The feedback process between DNA and the rest of the cell is species-specific; i.e., it is unique to each species and depends on the nature of the specific organism for its distinct activity.

---

[107] Denton, M., Kumaramanickavel, G., & Legge, M. (2013). Cells as irreducible wholes: The failure of mechanism and the possibility of an organicist revival. *Biology and Philosophy*, *28*, 31–52.

[108] For a comparison between our staunch hylomorphic view and a faint-hearted approach, along with a critique of both, see Hasker, W. (2013). The dialectic of soul and body. *American Catholic Philosophical Quarterly*, *87*(3), 495–509.

Second, more than DNA is passed on in reproduction. A single-celled zygote contains intricate machinery without which the DNA is biologically inert. This extra material is *always* co-present with DNA, and DNA requires the former for its specific functioning. Experiments have shown that “if a nucleus [which contains DNA] of one species is transplanted to the enucleated egg [an egg from which the nucleus has been removed] of an unrelated species, the egg may continue to develop for a while, *following the pattern characteristic of its own species, rather than the injected nucleus*—but the end result is premature death.”[110]

Can Organicism explain how changing a gene can alter characteristics of an organism? Consider an artist using a fine paintbrush to produce a painting. If an artist altered the brush, say one used to paint houses, and replaced it with a fine brush or even an ice pick, this change will alter his product. But neither an altered brush (or ice pick) nor its correspondingly altered product means there is no artist. Instead, the brush is a tool used by the artist; if something happens to either, the result will change.

Similarly, genes comprising DNA sequences are tools—instruments—and that is all. As H. F. Nijhout notes, certain genes produce (via interaction with other parts of the cell) certain materials, which, in turn, help determine which of various possible developmental pathways is actualized. According to Nijhout,

> Such genes can thus be said to control alternative developmental pathways, just as the steering wheel of a car controls the direction of travel. However, this is far from equating the steering wheel with the driver.[111]

What plays the role of driver? Brian Goodwin says the organism is an autonomous, irreducible center of activity, a whole with its own internal nature, its own species-specific principle of development in which the various parts are genuine functional entities that exist for and by means of each other and the whole of which they are parts.[112]

Clearly, this language implies that organisms are substances, not ordered aggregates or mere systems. If by “soul” we mean an individuated nature, the thin particular, then every living organism is identical to its soul, and it is plausible to take the soul to be what Goodwin is getting at when he talks about the organism as a whole.

Second, our particular version of hylomorphism—*Thin Particular Hylomorphism*—along with views (1) and (2) above, retain the soul or organism as a mereologically simple entity. Thus, these versions in particular ground the diachronic endurance of human persons through inseparable-part (e.g., physics entities, chemical entities, cells, organs, systems and processes) replacement. Moreover, we have suggested that whereas systems theory fails to ground diachronic endurance, all the versions of staunch hylomorphism are viable candidates for performing this metaphysical task, and for this and other reasons, they are superior to systems theory. Thus, systems theory does not provide an adequate defeater for our mereological argument, and we believe this is a solid reason for believing in a mereologically simple soul.

---

[109] Lewontin, R. (1992, May). The dream of the human genome. *New York Review of Books*, *28*. Cf. Noble, D. (2012). A theory of biological relativity: No privileged level of causation. *Interface Focus*, *2*, 55–64.

[110] Wells, J. (1993). Two dogmas of DNA. *Bible Science News, 31*(8), 13.

[111] Nijhout, H. F. (1990). Metaphors and the role of genes in development. *BioEssays*, *12*, 442.

[112] Goodwin, B. (1994). *How the leopard changed its spots*. New York, NY: Simon and Schuster.

This chapter completes our section on unity arguments for SD. In Chapter 9 we turn to a detailed examination of a specific version of a modal argument for SD.

PART IV

UPDATED AND NOVEL ARGUMENTS FROM MODALITY AND LIBERTARIAN FREEDOM

CHAPTER 8

Upgrading Modal Arguments for Substance Dualism

Philosophy is undergoing a resurgence of property and substance dualism.[1] One important argument for SD that has played a role in this resurgence is some version or other of a modal argument.[2] In what follows, we present and clarify one form of the argument, defend its premises, and respond to objections.

8.1 A MODAL ARGUMENT FOR SD

8.1.1 Statement of the Argument

The modal argument has its roots in Descartes (*Meditations* 2 and especially 6). Since then, several different modal arguments have been formulated, but we shall focus on one specific form. We are not implying that other forms are not appropriate; it's just not our purpose to survey various versions and instead, by presenting our own form of the argument, we can focus on what we take to be most important. Let SS be a wholly spiritual substance, e.g., a soul, and PO be a wholly physical object. Here is the argument:

1. The Indiscernibility of Identicals and $(x=y) \rightarrow \Box(x=y)$.
2. POs are essentially, wholly, and intrinsically physical and SSs are essentially, wholly, and intrinsically immaterial.

---

[1] See our discussion in Chapter 1.

[2] Moreland, J. P. (2014). *The soul: How we know it's real and why it matters*. Chicago: Moody Publishers.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

3. Possibly, I exist, and no POs exist.
4. My physical body is a PO.
5. Therefore, possibly, I exist without my physical body existing.
6. Therefore, it is not the case that I am essentially, wholly, and intrinsically my body or any PO.
7. I am essentially, wholly, and intrinsically either a PO or SS.
8. Therefore, I am essentially, wholly, and intrinsically an SS.

8.1.2 Clarification of the Argument

Certain clarifications are in order. In (1), “x” and “y” take Kripkean Rigid Designators or Swinburnean Informative Rigid Designators as substitutions. Thus, (1) expresses *de re* and not *de dicto* reference and modality. We will take up a defense of the essentialist part of (2) below, but here we want to clarify our use of “intrinsically.” This is mean to capture the idea that what makes a PO physical are the physical natures of the constituents—e.g., properties, property-instances, relations, relation-instances, parts, stuff, or events/processes—that “make up” a PO, that are inherent in its being. With proper substitution, the same goes for SS. Finally, we take “is a constituent of” to be unanalyzable and primitive.

In (3), (4), and following, “I,” “PO,” and “my physical body” are *de re* rigid designators.[3] We don’t like using “my physical body,” because “my” is an indexical possessive adjective, and we agree with those who take the nominative pronoun “I” and the possessive adjective “my” to be (pure) indexicals. As such, they are unique kinds of rigid designators: singular referring terms that refer to *sui generis*, irreducible nonphysical (in these two cases, mental) entities.[4] We developed this point in Chapter 6. Thus, designating something as “my body” is already to adopt a mental ontology as an “aspect” of the referent. But we set this issue aside, and for convenience, and in conformity with standard statements of modal arguments for SD, we shall retain “my physical body.” Finally, along with many philosophers, we take terms such as “I” and “my” to be implements of direct reference and, when used to express beliefs, e.g., that I am tired, they express *de re* (sometimes called *de se*) beliefs.

Regarding (7), if we limit our domain of discourse to strictly physicalist views and various dualist views of the subject of consciousness, the I, and of the physical body, then (7) is an exhaustive dilemma. If we remove such a limitation, (7) still captures the live options for most who support or reject the modal argument(s), so we set aside other options, e.g., contemporary idealism and panpsychism, for convenience. Let us grant (7) and our limited focus for the sake of argument.

We are assuming (1), (4), and (7). (5) follows from (3) and (4), (6) follows from (2) and (5), and (8) follows from (6) and (7). (2) and (3) are the main premises subject to attack. In what follows, we defend (2), (3) and respond to general defeaters of the modal argument.

---

[3] For a treatment of the importance of recognizing that these are *de re* and not *de dicto* expressions, see Meixner, U. (2016). Against physicalism. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 17–34). London: Routledge. See especially 24–25.

[4] See Madell, G. (1981). *The identity of the self.* Edinburgh: Edinburgh University Press; *The essence of the self.* (2015). New York, NY: Routledge.

8.2 DEFENSE OF THE ARGUMENT

8.2.1 Contingent Physicalism and Premise (2)

The vast majority of philosophers would embrace the notion that wholly physical and wholly mental particulars are essentially such. There are no possible worlds, W_1 and W_2, such that an object that is a PO in W_1 is? wholly nonphysical in W_2, and similarly with an object that is an SS. Speaking of whether physical objects, e.g., a bronze statue, could exist without being physical, Jaegwon Kim asserts what most take to be obvious: "The answer seems a clear no. If anything is a material object, being material is part of its essential nature; it cannot exist without being a material object."[5]

8.2.1.1 Merricks on Contingent Physicalism and the Modal Argument

But Trenton Merricks demurs, advocating contingent physicalism, which entails that an object could be wholly physical in the actual world but wholly immaterial in another world.[6]

> *Contingent Physicalism*: Given some object x and possible worlds W_1 and W_2, x is a PO in W_1 and x is a SS in $W_{2.}$

It follows from Continent Physicalism that the physicality of a PO is not essential to it. Merricks begins with a consideration of this argument:

1. For all *x* and all *y*, if *x* is identical to *y*, then *x* is necessarily identical to *y*.
2. Possibly, I exist and no physical thing exists.
3. I am identical with my body.

The dualist employs (1) and (2) to establish the negation of (3). Merricks seeks to block the inference from (2) to the negation of (3), and he claims that inference rests on:

4. M (my body) is essentially a physical thing.

Unfortunately, says Merricks, there is no argument for (4). Now consider:

5. Possibly, M (my body) exists and no physical thing exists.

Premise (5) entails the denial of (4) and deciding which to adopt depends on arguments/judgments regarding the acceptance of (3) or (4). The contingent physicalist will prefer (3) and also accept (1) and (2). From the fact that there is a possible world in which I exist and no physical thing exists, it does not follow that in the actual world, I am not identical to my body.

But what if someone argues that (4) is intuitively evident and properly basic? Merricks responds that (4) actually rests on:

6. It is necessarily true that a body is a physical thing.

---

[5] Kim, J. (2011). *Philosophy of mind* (3d. ed., p. 40). Boulder, CO: Westview Press.

[6] Merricks, T. (1994). A new objection to *a priori* arguments for dualism. *American Philosophical Quarterly*, *31*(1), 81–85; Cf. Zimmerman, D. (1991). Two Cartesian arguments for the simplicity of the soul. *American Philosophical Quarterly, 28*(3), 217–226.

But (4) does not follow from (7) as can be seen in this counterexample: It is necessarily true that the President of the United States is a member of the executive branch of the American government. But this does not entail that Bill Clinton is essentially a member of the executive branch.

Given that we have strong arguments/empirical evidence (e.g., split brain phenomena, mental/physical correlations,) and modal intuitions (e.g., we have strong modal intuitions that mental/physical interaction is impossible) against (4) and for (3) and given that we have only weak intuitions against (3), the physicalist should adopt (3) and give up (4), even though the modal argument for SD may persuade those half-convinced of SD or strengthen the faith of substance dualists. While this is a creative argument, we think it suffers from significant problems.

8.2.1.1.1 Objections to Merricks's Argument

First, a substance dualist could hold that (4) is properly basic and grounded in very strong modal seemings. For two reasons, this need not be a mere assertion. For one thing, as we will explain, Husserl has provided a detailed account of how we procure the relevant intuitions that provide access-internalist grounds for (4). The epistemic ground for (4) is not propositional knowledge, but a nonconceptual, (sub)-categorial intuitive awareness or apprehension of the features of formal ontology (material object) characteristic of one's body, as we will explain shortly.

Even if Husserl is wrong in his specifics, he shows that the proper-basicality claim need not be a mere assertion. Moreover, we will see shortly that dualist intuitions of disembodiment are historically and geographically ubiquitous, and they are naturally formed by little children with no exposure to dualist or religious teaching, much less to ancient Greek philosophers. A good explanation for this fact is that people have a direct awareness of their simple selves, a categorial intuition of their bodies *qua* being physical, and they become aware of their nonidentity and the contingency of their relation to each other. So, people do not have strong intuitions for (3). Quite the contrary. In fact, as argued in Chapter 5, many physicalists who believe (3) acknowledge that their intuitions against (3) remain.

One could respond that this shows merely that the folk ontology of human beings includes the contingent fact that we are possibly, but not essentially, disembodiable. However, as we showed in Chapter 3, if that were the case, one would expect to find a fair number of children and adults, *sans* prior physicalist indoctrination, expressing physicalist understandings of the afterlife. But that is not the case because people intuit the *kind* of thing they are—essentially immaterial.

Merricks counters that dualist commitment to (4) relies on (7):

(7) It is necessarily true that a (wholly physical) body is a physical thing.

We disagree and embrace a Husserlian account of (4) relying on the relevant intuitions. But more importantly, in context, (7) is a *de dicto* claim with wide scope for the modal operator, and “a body” seems to be a nonrigid designator. If this is a correct understanding of (7), then if there is a propositional ground for (4), it is not (7) but (7′):

(7′) My body is necessarily (more appropriately, essentially) a physical thing.

Here “my body” and “physical thing” are rigid designators, the premise involves *de re* necessity with the modal operator exhibiting narrow scope. So Merricks's employment of (7) is not relevant.

One could respond to our argument by advancing a form of descriptivism, roughly, the view that mental contents (or linguistic entities) "directed toward" individuals are definite descriptions that uniquely pick out those individuals that satisfy them. Thus, there is no such thing as direct, *de re* reference. After all, one cannot think about something, the response continues, unless one thinks of it as propertied in one way or another, and the descriptive mental content expresses that unique way the referent is so propertied.

In reply, Walter Hopp notes that this objection is based on a confused understanding of what direct, *de re* reference does and does not entail.[7] For one thing, it is a false dichotomy to hold that directly referential terms, e.g., proper names, natural kind terms or categorial expressions, either have descriptive meaning (and, therefore, are not directly referential) or no sense at all. Instead, some such terms have descriptive and directly referential nondescriptive aspects. While one cannot entertain demonstrative thoughts about a chair directly in front of one without being aware of at least some of its features, this does not entail that there is no nondescriptive, directly referential element inherent in true demonstrative reference.

Moreover, just because one must think of, say, New York, in a certain way to think about it, it doesn't follow that one thinks of New York via a uniquely satisfied definite description, viz., as just whatever properties one thinks New York to have. One can directly refer to New York *and* keep it in mind as one varies the properties one thinks New York to have. As Hopp notes: "Ultimately, the difference between thinking of something directly and thinking of it as just whatever has such and such properties is phenomenologically evident. But thinking about something directly is not incompatible with thinking of it in some way or via some sense."[8]

Edmund Husserl clarified this claim quite precisely.[9] Husserl began his account of intentionality and *de re* reference by bracketing (a.k.a. *epoche*) the natural attitude in favor of the phenomenological attitude (a.k.a. *phenomenological/transcendental reduction*). This meant that he set aside scientific considerations about the existence and nature of the physical world due to his conviction that when it comes to the realm of consciousness, "Natural cognition begins with experience and remains *within* experience."[10] For Husserl, the phenomenological exploration of consciousness is a strictly first-person endeavor.

In his account of intentionality, he held that every intentional experience is constituted by two independently variable yet inseparable moments—a matter and a quality. The act-quality is roughly what we mean today by the attitude in propositional attitudes. The quality is what grounds the intentional act as a judging, hoping, merely experiential, etc., one.

More important is Husserl's assay of an intentional experience's matter. The matter gives the intentional act its directedness toward an object and consists in a referring to the object in a specific manner. It fixes reference on this very object in this very manner. Put differently, the matter included *a reference to* an intentional object and the manner in which it is intended, what it is *intended as*. The former stands to the latter as determinable to determinate. The matter presents the object in a specific way. The quality and matter of an intentional act are the act's intentional essence, and in purely intuitive acts, e.g., perceptual, eidetic, or categorial, the intentional essence is entirely nonconceptual. In such acts, the nondescriptive, directly referential elements secure *de re* reference.

---

[7] Hopp, W. (2011). *Perception and knowledge* (p. 21). Cambridge: Cambridge University Press. Cf. 12–26.

[8] Ibid., pp. 21–22.

[9] Husserl, E. (1970). *Logical investigations* (2nd ed., J. N. Findlay, Trans.). London: Routledge and Kegan Paul (hereafter, *LI*), section 17: 578–section 22: 593 (especially section 20).

[10] Husserl, E. (1982). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy. First book: General introduction to a pure phenomenology*, translated by F. Kersten (The Hague: Nijhoff, 1982), I section 1: 5.

Finally, Husserl held that intentionality is a monadic property and not a relation. Among other things, this permits him to avoid hunting around for a surrogate object (one's own concept, a haecceity, and so forth) in cases where one is aware of or thinking about a nonexistent object like Zeus, a surrogate that is clearly *not* the intended intentional object.

It may be useful to compare Husserl's view of direct *de re* awareness and reference to an intentional object with a sophisticated alternative presented by John Campbell. In his major work *Reference and Consciousness*, published in 2002, Campbell focused his study on the reference of perceptual demonstratives to currently perceived objects on the basis of current perception of them.[11] Knowledge of reference of perceptual demonstratives is grounded in the basic idea of direct acquaintance with an object in terms of nonconceptual direct awareness of the object.[12] Such awareness involves conscious attention, a primitive psychological relation that provides us with knowledge of reference. This also involves a primary experience of location, with a subsequent Gestalt organization binding together the referent's properties, facilitated by unconscious sub-mechanisms described by science that stream various forms of sensory information.

In his 2002 publication, Campbell adopted a relational view of experience, viz., as a simple relation between a subject and an object.[13] The qualitative character of a relevant experience is constituted by the qualitative character of the scene perceived. In this way, perceptual experiences (allegedly) put us in direct contact with the world by being partly constituted by external objects. In a 2009 publication, Campbell tackles the problem of accounting for the fact that we can be consciously aware of an object from different standpoints. In response, he expands his analysis of conscious awareness such that it is a three-placed relation among a person, a standpoint, and an object.[14]

Campbell agrees with Husserl that words, e.g., demonstratives and *de re* referring terms, have derived intentionality with the relevant mental acts or states possessing fundamental intentionality. They also agree that *de re* reference is experientially, conceptually, and epistemically fundamental relative to *de dicto* counterparts, and that in typical cases of (especially demonstrative) reference, the thing itself must be directly given.

However, in our view, Husserl's ontological categories and distinctions are far richer than those of Campbell, his adoption of monadic over relational accounts of intentionality and awareness is preferable to Campbell's views, and we agree with Husserl's bracketing, something clearly eschewed by Campbell. Regardless of the reader's agreement with these last points, the items shared by both thinkers will be crucial in our discussion of the modal argument in what follows. And they support the claims that (i) the ground for (4) (M (my body) is essentially a physical thing.) is directly experiential and (ii) the previous claim in (i) is no mere assertion and has been supported by various well-developed theories.

Hopp also observes that since Sellar's attack on the "myth of the given," there has been a widespread consensus that reasons-giving relations hold among only mental states whose intentional contents are conceptual or propositional. Thus, perceptual (or nonperceptual intuitive states) must include conceptual/propositional content.[15] Hopp calls this experiential conceptualism, and while it may not entail the reduction of *de re* to *de dicto* reference, minimally, experiential conceptualism is hard to square with direct *de re* reference and our epistemically justified confidence in its success but fits quite nicely with *de dicto* reference.[16]

---

[11] Campbell, J. (2002). *Reference and consciousness*. Oxford: Clarendon. While his account is limited to perceptual demonstratives, he does include acquaintance with properties and, in any case, his views could easily be expanded to any object known by direct intuitive awareness of some sort or another.

[12] Ibid., pp. 1–2; chapters 1–2, 5–7.

[13] Ibid., chapters 6–7.

[14] Campbell, J. (2009). Consciousness and reference. In B. P. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford handbook of philosophy of mind* (pp. 657–659, 662). Oxford: Clarendon.

Given what seems to be the phenomenologically accurate depiction of fulfillment in Husserl (Section 8.2.2.2.3), all *de dicto* intentionality is grounded in *de re* intentionality, since the latter always involves a genuine, direct reference to some specific individual, property, relation, etc., such that the fulfillment of the *de dicto* expression requires that referent "to show up in person," being directly present or given to the referring subject in the fulfilling act. In this way, purely nonconceptual, intuitive—e.g., perceptual—intentional states are reason-giving with respect to the relevant conceptual/propositional content in that the former are the fulfillment of the latter.[17]

In contrast to experiential conceptualism, this view makes *de re* reference the foundation of *de dicto* reference. However, ubiquitously favored forms of experiential conceptualism entail that there is no direct intuitive awareness of an intentional object "out there" and, without this, *de re* reference becomes otiose. And since experiential conceptualism also entails that the only reason-giving relations that obtain are conceptual and propositional, this would seem to favor quite strongly a commitment to irreducible *de dicto* reference and an eschewal of *de re* reference as epistemically irrelevant.

These remarks about experiential conceptualism and Husserlian fulfillment structures may not be factors in Merricks's preference for *de dicto* treatments of the relevant referring expressions in our modal argument. Nevertheless, it is still true that many philosophers who embrace experiential conceptualism would favor *de dicto* interpretations of these expressions over our promotion of them in *de re* terms.

For at least these philosophers, a Merricks-style critique of our use of (4) and ($7'$) in favor of (7) will come down to the debate between experiential conceptualism and Husserlian-type understandings of fulfillment structures or some other experiential internalist alternative to them.

Second, in criticizing van Inwagen's simulacrum hypothesis, Merricks says that even if there is a possible world in which van Inwagen's view is true, nevertheless, it is the actual world that is of concern to us and his view is false in the actual world.[18] Analogously, he claims that (2) of his argument implies there is a possible world in which I am a wholly immaterial object, but that world is not the actual one.

We offer two responses. First, modal intuitions, especially as understood by Husserl, are of *kinds* of objects—the kind "pure ego" and the sub-categorial kind "physical object"

---

[15] Hopp, pp. 1–2.

[16] There is a widespread tendency on the part of many naturalists to reduce or replaced *de re* to *de dicto* reference. For a classic example of this, see Loar, B. (1997). Phenomenal states. In N. Block, O. Flanagan, & G. Guzeldere (Eds.), *The nature of consciousness* (pp. 597–616). Cambridge, MA: MIT Press. For an excellent response, see Robinson, H. (2016). *From the knowledge argument to mental substance* (pp. 73–92). Cambridge: Cambridge University Press.

[17] For Husserl, pure perceptual states have two kinds of nonconceptual content—intuitive content (the object is directly present) and horizontal content (roughly, one is now aware that the physical object of which one is directly aware has more to it and one is indeterminately aware of this more while anticipating further directly presentational experiences of this "more.)"

[18] Merricks, T. (2001). How to live forever without saving your soul. In K. Corcoran (Ed.), *Soul, body and survival* (pp. 183–200). Ithaca, NY: Cornell University Press. See 84.

(including “wholly physical body”). Merricks could reject this as irrelevant, but he shoulders a burden of proof here since most philosophers and lay people find contingent physicalism to be unacceptable. David Lewis is not the only one who must face the incredulous stare. Merricks also needs to interact with modal epistemologies like Husserl’s.

Second, we think Merricks’s notion that a wholly physical body could exist in a possible world as a wholly immaterial entity is, in part, the result of taking possible-worlds talk as a useful semantics for modal operators and, in some way, ontologizing possible worlds. We can’t argue the point here, but if something like Alexander Pruss’s Aristotelian-Leibnizian view is correct, then it entails that modal truths are grounded by the causal/dispositional capacities of actual substances.[19] Assuming this is correct, if an alleged “wholly physical body” in the actual world can exist wholly immaterially, then the former must have a wide-range of immaterial capacities in the actual world and, arguably, this would entail that it is not *wholly* physical in the actual world.

Third, it seems unintelligible how complex objects x and y could be identical (in this or across worlds) if they share no proper constituents. How could the angel Gabriel in one world be a frog’s body in another when they share no proper constituents? To clarify, a complex object is one with two or more proper constituents. A proper constituent of some object x is an entity that would be included in an ontological assay/inventory of all and only those entities that are “within/enter into the being of” or that “make up” x. “Proper constituent” includes parts, properties/property-instances, intrinsic relations/relation-instances, structure/structure-instances, and events/processes.

For Merricks, one’s body in the actual world is a PO and in another world an SS/IO. There is nothing shared between a PO and an SS/IO that could ground or account for their identity.

Look at it another way. If a PO is wholly physical contingently, then physicality is accidental to it, e.g., to one’s body. But if a proper constituent (or structural arrangement of such) is accidental to an entity e, then e could exist without it. Drawing insights from other relevant cases, we conclude: The object *has* the accidental proper constituent(s) in the actual world and, thus, in another world where that proper constituent is lacking, the object exists yet lacks the accidental proper constituent. Thus, the object is what *has* and is *not identical to* that constituent. The same would be true of an accidental proper-constituent cluster. So, the object must be something “over and above” that cluster, it has different clusters in different worlds, and it retains identity in both worlds. The object’s identity is independent of its contingent features since *it* can exist without those features.

Applied to one’s body, if it is to retain identity in different worlds (and not merely involve counterparts), then there is more to one’s body than physicality or immateriality. It cannot be a PO or SS. One’s body *has* physicality and then immateriality accidentally. But one’s body must have its essential identity located in something additional to these accidental aspects in order to be the same entity in each of the relevant possible worlds to make sense of what remains the same and has one and now the other accidental proper-constituent clusters.

---

[19] Pruss, A. R. (2011). *Actuality, possibility and worlds.* London: Bloomsbury, Parts V, VI. Pruss would include the Leibnizian component of God being one of the relevant substances. For an alternative dispositional account of possible worlds talk, see Martin, C. B. (2007). *The mind in nature* (pp. 1–4, 29–33). Oxford: Clarendon Press.

Admittedly, Merricks is not arguing that a wholly physical object could *become* a wholly immaterial object. Still, an analogy with Aristotelian accidental change is appropriate and illuminating. Consider a dog changing from being black to gray. For this to be change and not successive replacement involving ceasing-to-be and coming-to-be, there must be some entity that is self-identical throughout the change and that had blackness and now has grayness. That entity is "over and above" the color-properties.

How might a contingent physicalist respond? We think the best line would be to say that identity, including the identity of a complex object, is brute and ungrounded; it is the object itself—a PO/SS—that is identical in the two worlds, and that's that.

Unfortunately, this claim comes very close to begging the question. We are disputing that a PO could be identical to an SS as Merricks depicts it, and this response says, in effect, that our arguments are ineffective because they just are identical and there's nothing more to be said. Surely, when it comes to complex objects as we have characterized them, we are owed at least some sort of account as to how such an identity could be brute. Not all brute primitives are created equal. We could see how a constituentless bare particular's self-identity could be brute, but complex objects are another matter altogether.

Further, we think the following claim is ambiguous: Identity is never grounded ("each thing is what it is and not something else" is about all one can say on the matter), including the identity of complexes, so the search for a ground is misguided. If this means that the identity relation is primitive and irreducible to some other relation, e.g., exact similarity or some (weaker) persistence relation, then we agree. But we suspect that this isn't what is being affirmed. We suspect it means that there is no reductive analysis of cases of identity themselves; in these cases, identity is what it is and, consequently, uninformative.

A good place to begin evaluating this interpretation is the debate over diachronic personal identity between advocates of the simple and complex views.[20] One criticism of the simple view is that by taking personal identity to be primitive and brute, the simple view is uninformative and merely an assertion of diachronic identity.

This criticism is confused in that it conflates a reductive analysis of diachronic identity with an ontological assay of the entities that retain it. The simple view does entail that no reductive analysis of personal identity, e.g., in nonpersonal terms involving sortal-dependent psychological or bodily persistence conditions, is adequate. But this does not entail that a case of diachronic identity is uninformative because in each case—particularly, cases of wholes that are metaphysically complex and mereologically simple with respect to separable parts—an ontological assay may be given of the entity that retains identity to understand what must be the case for identity to obtain.

For example, a simple advocate could say that personal identity is identity of a specific type of soul and, moreover, that the identity of that sort of thing requires identity of its constituents—the essence "human personhood," the nexus of exemplification and primitive thisness, e.g., a bare particular. Such an assay retains the basicality of identity itself but provides a nonreductive ontological inventory of the relevant complex entity and insight as to what must stay the same for the entity to retain identity. In a property-exemplification view of event identity,

---

[20] See Gasser, G., & Stefan, M. (Eds.). (2012). *Personal identity: Complex or simple*. Cambridge: Cambridge University Press.

> *Property-Exemplification Event Identity*: $(e_1=e_2)\rightarrow[(S_1=S_2) \wedge (P_1=P_2) \wedge (t_1=t_2)]$, where e_1, e_2 are events, S_1, S_2 are substances, P_1, P_2, are properties, and t_1, t_2 are times.

Here, identity itself is primitive, but the natures of the entities (events) that are/are not identical are complex and in need of an ontological assay. We suggest this same response is sufficient to defeat this objection.

Finally, it may be objected that parthood is constituted by sameness of causal-functional role integration with respect to a whole or by being caught up into the same life. So, if the parts themselves are replaced by other types of parts that play the same role or are caught up into the same life, identity is retained.

In response, many philosophers, e.g., Jaegwon Kim, take functional roles as concepts, as ways of taking something, and not real entities that constitute the relevant objects. *Sans* some sort of Aristotelianism, we agree.

Second, this confuses an ontological account of what makes something a part with the nature of the part itself. The latter is essential for macro-identity. Thus, if all the parts of an object are replaced with different parts that play the same role, you have different objects that are functionally isomorphic, rather than the same object.[21]

Third, it is far from clear that a functional role construed ontologically is a physical entity and, if not, then such a PO is not wholly physical. Finally, being caught up in the same life faces this problem: if the activities of the parts of O_1 at t_1 are different from those of the parts of O_2 at t_2, how is it the same life? Since a life is an aggregate of living activities, it seems that if those activities change, so does the life. Thus, life is not stable enough to ground an organism's identity. And if all the physical parts are gradually replaced with immaterial parts and all parts exhibit life-constituting activities, then the life at t_1 will be a set of physical activities, the life at t_2 a set of mental activities, and the life at one time is just not identical to the life at another.[22]

We want to tweak this last point to support our second objection above. Suppose O_1 at $t_1 = O_2$ at t_2 because their parts' activities constitute the same life. Assuming O_1 is a physical object composed of and only of physical parts that exhibit physical activities, then O_2 seems to be a physical object. But suppose, instead, that O_2 at t_2 is composed of and only of purely immaterial animal spirits exhibiting physical activities. If we still assume that O_1 at t_1 is entirely physical and that the activities of the parts of O_1 and O_2 constitute the same life, then, based on our criterion of parthood, it would seem that O_2 is a physical object as well. But we have strong intuitions that resist this conclusion. O_2 is composed of and only of immaterial parts with no physical constituents. Given this, surely O_2 is an immaterial object and not identical to O_1. But then, the nature of an object's constituents determines its ontological kind, and criteria for parthood simply tell us which parts are relevant in deciding the ontology of the whole.

To close out our response to Merricks, (1) we think that our modal intuitions for a genuine metaphysical distinction between the soul and body are much stronger than those against mental/physical interaction (note the ubiquitous dualist intuitions cited above and the ubiquitous intuitions that a wholly immaterial God could cause things to occur in a wholly physical world); (2) strict physicalism, mere property dualism, and substance dualism are empirically equivalent and, thus, there is *no* empirical evidence that favors one of these three view (we invite the physicalist to present such empirical evidence that a substance dualism could not easily accommodate); and (3) we have tried to show that there are good arguments against contingent physicalism and the arguments for it are weaker than Merricks supposes. Thus, we see no epistemic considerations here that favor physicalism over substance dualism.

---

[21] This claim does not apply to genuine substances (with no separable parts) with substantial forms and real, irreducible teleology.

[22] We are indebted to Andrew Bailey and Joshua Rasmussen for their helpful input in this discussion of Merricks's view.

8.2.1.2 Contra Bailey's Contingent Physicalism

8.2.1.2.1 Four Features of Bailey's Position

Andrew Bailey is another advocate of contingent physicalism, and he offers a very interesting, unique approach to defending it.[23] Here are four main features of Bailey's view.

1. Bailey on the Categories: In developing his set of ontological categories, Bailey locates himself squarely within "the grand tradition of ontological categorizing" (Bailey 20) and is "doing old-fashioned categorial ontologizing" (Bailey 21). Accordingly, he seeks a realist set of categories that specifies "at the highest level of abstraction how reality divides" (Bailey 5). The categories he employs are these: Level I: Thing; Level II: Abstract Property and Concrete Object; Level III: branching out of Concrete Object are Material Object and Immaterial Object. Level IV: branching out of Material Object are Thinking Material Object and Unthinking Material Object; branching out of Immaterial Object are Thin Immaterial Object and Unthinking Immaterial Object. Finally, a curious feature of Bailey's categories is that in at least some cases, members belong to their category contingently such that they could change categories. Specifically, items in Thinking Material Object in one world belong to Thinking Immaterial Object in another world.

   Setting aside the categorial box containing God, the chart itself is shown in Figure 8.1.
2. Bailey's characterization of Wholly Material and Wholly Immaterial Objects (see Figure 8.1): While acknowledging that he lacks "rigorous definitions of the key ideas here—mentality, material object, human person, spirit" —he offers remarks about them he takes to be adequate for his purposes. For Bailey, x is a wholly material object *iff* at some level or other *x* is composed of items (separable parts?) with narrow physical properties (those figuring into fundamental physics) none of which thinks (Bailey 10–11). As far as we can tell, Bailey almost entirely characterizes a spirit by way of the *via negative*, viz., as an incorporeal or immaterial object.

However, in one place he says that spirits are such that at every level they have parts (or are themselves things) that either think or are not characterized by fundamental physics. Since it seems that the only kind of part with which Bailey is familiar is a separable part, by granting the possibility that a spirit has parts, indeed, thinking parts, his view

---

[23] Bailey, A. (2021). *Monotheism and human nature*. Cambridge: Cambridge University Press. Subsequent page numbers within the text refer to this book. Bailey is a Christian theist. Accordingly, he would most likely not ascribe contingent physicalism to God.

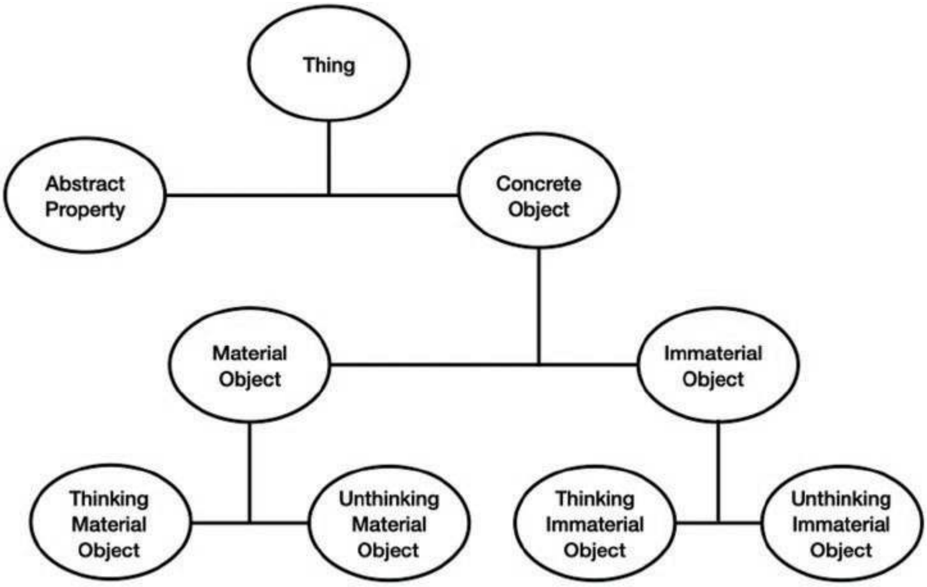


FIGURE 8.1 Bailey's Ontology.

is at odds with most dualists, including us, about the simplicity of a spirit.[24] In any case, we don't know how important this depiction is because in virtually all other cases, he uses "immaterial" or "incorporeal" thing. In fact, on the same page as the depiction just mentioned, Bailey returns to his usual pattern and tells the reader that from that point on, he will use "wholly immaterial" or "wholly incorporeal" thing for what it is to be a spirit.

3. Bailey's Thought Experiment in defense of *Contingent Physicalism*: According to Bailey, our nature is highly contingent (Bailey 9) and, more specifically, in the actual world we are wholly material objects, yet there is a possible world in which we are wholly immaterial objects. He invites us to consider the following thought experiment (Bailey 14): Imagine that Bradley, a living organism, has at some very low level a multitude of tiny physical parts we'll call atoms. Through a complex network of integrated and united causal dispositions, the atoms compose Bradley. In general, a sufficient condition for some atom *x* to be a part of a living organism O is for *x* to play a functional role in O's complex network of causal relations.

One day, God annihilates one of Bradley's atoms and replaces it with an angelic surrogate that is functionally equivalent with the replaced atom and, thus, exhibits the integration and unity to count as a part of Bradley. God continually reiterates this process, atom by atom, angel by angel, until after a few months, the transformation is complete. Since angels are wholly immaterial objects, Bradley is now at some low level, entirely composed of immaterial parts, and thus, Bradley is a wholly immaterial object. Throughout the process, Bradley sustains absolute Leibnizian identity.

---

[24] In this regard, Bailey's view is a close cousin to classic Mormon anthropology. See Moreland, J. P. (2002). The absurdities of Mormon materialism: A reply to the neglected Orson Pratt. In F. Beckwith, C. Mosser, & P. Owen (Eds.), *The new Mormon challenge* (pp. 243–270). Grand Rapids, MI: Zondervan.

4. Bailey's Assumptions: Bailey lays out four assumptions on which his argument rests (Bailey 14). Three of them are relevant to our discussion.
    (i) Causal interaction makes for parthood. If something assumes an atom's role in a complex network of causal relations, it becomes part of the greater whole just as the atom was.
    (ii) If something is, at some very low level, composed entirely of wholly immaterial thinking beings, then it is itself wholly immaterial.
    (iii) Something can survive the replacement of one part with a surrogate that exactly duplicates that part's causal role.

8.2.1.2.2 Four Problems with Bailey's Position

What should we make of Bailey's argument? It is surely a fun, creative, and interesting bit of philosophy. But in our view, there are several problems with it and, in the end, we judge it a failure to establish the truth or plausibility of contingent physicalism. Many of the objections raised above against Merricks's view apply with equal force to Bailey's contingent physicalism. We will not repeat them unless a unique aspect of Bailey's views presents an occasion to restate one of these arguments from a slightly different perspective. Further, while we want to gather into one place the general problems with Bailey's position, some of our points are developed more fully elsewhere in this book. In those cases, our remarks will be brief, and we will direct the reader to other parts of the book where a fuller presentation is found. With that said, here are four general problems with Bailey's position, some of which have a few distinct subpoints.

First, Bailey's depiction of living organisms, especially human persons, is question-begging and, most likely, false in the actual world. Bailey appears to be presenting a view of living organisms as they are in the actual world. We say this for two reasons. First, he claims that his categories, especially those with human persons as items, present truths about us (Bailey 5), specifically, contingent truths that are actually true of us even though they could have been false (Bailey 6). And by pointing to a node, you say "that's our place in this world" (Bailey 6).

Second, Bailey begs the question against many substance dualists and staunch hylomorphists. Bailey's thought experiment is intended to show that a wholly material being could become a wholly immaterial being. Since he is a physicalist—specifically, an animalist of some sort—he holds that we are actually wholly material objects. And while his thought experiment is diachronic in that a wholly material object *becomes* a wholly immaterial object, given its speculative nature, we think the real point of the experiment, especially as a defeater of a modal argument for substance dualism, is that just because there is a possible world in which we are wholly immaterial, it does not follow that in the actual world we are not wholly material.

If we are right about this, then Bailey begs the question against many dualists like us in his depiction of the soul or of biological organisms. Many (especially Cartesian) substance dualists do not grant that we are identical to living organisms. Rather, we are simple souls. Moreover, staunch hylomorphic dualists do not think that living organisms are composed of separable parts; rather, they are substances that have no such parts and, thus, a compositional, part-whole approach to organisms is wrong (see Chapter 7 for a presentation and defense of versions of staunch hylomorphism).

In fact, Bailey's depiction of living organisms appears to be an expression of some version of animalism or a close cousin (e.g., a modified form of constitutionalism), and we have evaluated animalism in Chapter 6. Bailey does not so much as mention in a footnote that there are problems with his depiction and alternatives to it. Thus, his appropriation of that depiction is question-begging and far too quick. We need an argument here, especially since his thought experiment and other relevant claims are presented as capturing the truth in the actual world.

Further, we argue elsewhere and provide significant documentation that well-developed philosophical versions of staunch hylomorphism and scientific versions of organicism have greater explanatory power than does the depiction of organisms presented by Bailey and, thus, the latter is most likely false (Bailey 62–77).[25]

Second, Bailey's living organisms are mereological aggregates (MAs). Accordingly, they inherit a number of successful defeaters for such aggregates and for his appropriation of them in building his case. In Chapter 7, we presented a thorough depiction of an MA, and we pointed out a number of problems with wholes construed as such. Recall that, setting aside eliminativism, an MA is a weakly unified collection of separable parts that stand in external relations to other such parts; they and their parts exhibit efficient causality and play no functional role (nor exemplify irreducible teleology) which is intrinsic to the wholes or their parts, and are depicted, at least in the actual world, as exhibiting part-priority, usually, micro-physical part priority. By "weakly unified" we mean (i) synchronically, the parts exist and have their identity independently from the whole of which they are parts and, thus, the whole is merely a collection of parts standing in certain external (e.g., causal) relations. On a sliding scale from mere heaps (e.g., a heap of sand) to classic Aristotelian substances, mereological aggregates fall toward the heap end of the scale; (ii) diachronically, MAs are subject to mereological essentialism constraints regarding endurance and perdurance.

We take Bailey's characterization of living organisms to render them MAs, though we do not believe that was his intent. He would most likely reject our view on this. However, we claim that they are MAs because (i) Bradley is his paradigm case, and it is clear that from First-Bradley (who is wholly material) to Final-Bradley (who is wholly immaterial) the collection of parts throughout the transformation involve separable ones; (ii) the causal relations must be external if the relata (the separable parts) retain their identity whether or not they are in the whole or standing in these relations; (iii) Bailey claims that the parts play a functional role, but his metaphysics of living organisms disallows them to do so literally. MAs are like artifacts: They (or their parts) may be described with functional concepts, but these are ways of *taking* the relevant entities, of describing them *as if* they exhibited functionality. The employment of such concepts does not literally ascribe intrinsic functional roles or teleology to the MAs or their parts. Remember, the relevant "complex network of causal relations" is constituted by and only by efficient causal relations. No part actually plays a role in such a network.

If this is correct, then Bailey's employment of his criterion for parthood (Causal interaction makes for parthood. If something assumes an atom's role in a complex network of

---

[25] Our use of "organicism" in these pages is quite different from Bailey's usage. For him, "organicism" seems to be equivalent to "animalism." See Bailey, A. M. (2016). Composition and the cases. *Inquiry*, *5*, 453–470.

(efficient) causal relations, it becomes part of the greater whole just as the atom was.) is not possible since there are no functional roles the parts literally play.

And even if somehow Bradley smuggles in literal, intrinsic functional roles/intrinsic teleology, not only would this appear to be an odd *ad hoc* move, it would also render it difficult to claim that living organisms *qua* MAs (or animals) are wholly material objects. We know of no one who treats intrinsic teleology as physical, and consequently, physicalists avoid attributing such to wholly material objects, preferring instead to give a physically acceptable reduction (e.g., etiological) to teleology. Moreover, the paradigm cases of functional roles are cultural artifacts (being a quarterback, a dollar bill, etc.) that are constituted as such by the intention(s) of the artifact's inventor or by the culture in which the artifact is located and identified. Thus, it is hard to see how a functional role is physical.

If one responded by depicting a functional role in terms of intrinsic teleology, then such teleology, and the metaphysical commitments within which teleology makes the most sense (e.g., substantial forms, inseparable parts) imply that living organisms are not, in fact, wholly material objects. For more on the points as they apply to animalism, see Chapter 6.

Third, several aspects of Bailey's thought experiment render it a failure. We offer four such aspects. To begin with, Bradley is a mereological aggregate and, as such is subject to mereological constraints. We have treated this issue in considerable detail in Chapter 8, including the idea that somehow sameness of structure or "causal network" can ground endurance through part replacement. Thus, from the first part replacement to the wholly immaterial Final-Bradley, we have a succession of different wholes which are not identical to the wholly material First-Bradley. So the thought experiment is otiose.

Perhaps Bailey could object by appealing to animalism as a better depiction of Bradley. So far as we can tell, Bailey has never clarified the necessary and sufficient conditions for a differently composed organism at different times to be a genuine continuant. In any case, as we noted above, if Bailey grounds the fact that Bradley is a continuant on the fact that at each stage of Bradley's transition, the activities of whatever parts are caught up in the same life, then this move fails. Life itself is in flux through time as the set of living activities is constantly changing. Thus, that set is itself an MA void of the enduring stability needed to ground Bailey as a continuant.

In addition, his thought experiment makes clear why one of our objections above to Merricks's contingent physicalism has force. Recall that we objected to Merricks's claim that a wholly material object in this world could be identical to a wholly immaterial object in some possible world because the two objects have nothing whatsoever in common to ground or make sense of their identity. Rather than the two objects being exhaustively (wholly) constituted by their material or immaterial constituents, if "they" retain identity, it seems that their physicality and materiality are possessed by something contingently, and that something is what remains the same. It *has* physicality in one world and immateriality in another. We also provided defeaters for the claim that identity in this case is brute.

Bailey's thought experiment makes explicit why this is a problem. Remember, First-Bradley is a wholly material object. How so? Well, given Bailey's compositional approach to analyzing Bradley, all of his separable parts are physical as are all the external causal relations exemplified between and among them. We may add physical properties that characterize these parts and relations, and, perhaps, the whole they compose, and physical (efficiently causal) activities of the parts, as long as everything employed in the analysis is physical. Besides these constituents, there doesn't seem to be anything else to Bradley. Now think of Final-Bradley, a wholly immaterial object. Every constituent of Final-Bradley must be immaterial, e.g., his angelic parts, and the exertion of active power to push and pull other angels.[26]

Now ask: What possible entity do First- and Last-Bradley have in common that either grounds their identity or is the item that is just identical in the two stages? There is no answer to this question and, thus, no way of providing a metaphysical account of their identity. That identity is a mere assertion and, as noted, cannot be asserted as brute with propriety.

Granting that composition is not identity, one might object that the composite itself, e.g., the body as a wholly physical object, grounds identity through part replacement. There is nothing mysterious about this and, accordingly, no need for a further grounding explanation.

We responded to this kind of objection above, but we want to proffer additional considerations here. Limiting our focus to MAs and their inseparable parts, the dependence/grounding of the former in the latter is of two sorts: essentially$_1$ grounded and rigid existentially grounded:[27]

> *Essentially$_1$ Grounded*: (IG) x is essentially$_1$ grounded in $y =_{\text{def.}}$ there is a two-place predicate "F" such that the identity of x includes that x is related by F to y.
>
> *Rigid Existentially Grounded*: (RG) x is rigidly grounded in $y =_{\text{def.}}$ $(\mathrm{E}x \rightarrow \mathrm{E}y)$.

IG goes beyond existence grounding by specifying what is necessary for x to exist as the kind of thing it is. RG states that the existence of x depends on a specific entity, y.

Applied to MAs, x is the composite (e.g., the wholly physical body), y is all and only x's separable parts, and F is a predicate that expresses the part–whole relation. According to IG, if the ys add or lose a separable part, then x no longer stands in the part–whole relation to the ys. Call the new collection of all and only x's separable parts the z's. IG shows us why x loses its identity when the ys are replaced with the zs. Consequently, a composite cannot ground identity through part replacement because the composite itself does not endure.

If we are correct about all this, then while genuine Aristotelian substances which have (i) inseparable parts/modes (but are not composed of separable parts) are the sorts of things that can literally be constituted by intrinsic teleology and have (ii) inseparable parts essentially characterized by their intrinsic functional roles in the whole, Bailey's living organisms *qua* MA are capable of neither. They are like artifacts with *as if* and not real, intrinsic functional roles. This becomes clear when Bradley starts having more and more angelic parts. Note, that when an angel becomes a part of Bradley, it does not undergo substantial change and get its identity as an inseparable part from the teleological role it plays. Indeed, that role is not an intrinsic, constituting essence of the angel. Instead, the angel is *told what to do* and acts accordingly. The angel is not literally a functional entity whose role is part of its essence and identity. Rather, it functions *as if* it played a role.

---

[26] Bailey does qualify his views as follows: First-Bradley is a wholly material object in that at some very low level he is composed entirely of physical atoms. At higher levels, he still has cells and organs. We assume these are physical as well. Final-Bradley is wholly immaterial in that at some very low level he is composed entirely of wholly immaterial objects, viz., angels. But at higher levels, he retains cells and organs. We see how First-Bailey is wholly material since his cells and organs are wholly physical, being composed entirely of wholly physical atoms. But if Final-Bradley's cells and organs are like First-Bradley's, then Final-Bradley is not wholly immaterial, having physical cells and organs. Bailey recognized this problem and rightly pointed out that the cells and organs of Final-Bradley are wholly immaterial, being entirely composed of wholly immaterial parts (angels). See Bailey (2021) p. 16.

[27] See Inman, R. (2019). *Substance and the fundamentality of the familiar* (pp. 62–74). New York, NY: Routledge.

This is importantly analogous to the way the parts of an artifact "function" in a whole, e.g., the parts of a car's engine. The engine's parts engage in mere efficient causation and do not literally play any functional role. But the car's inventor designed the parts and by doing so, "told the parts" as it were, to behave via efficient causation in a way that an observer could act *as if* the part played a functional role. Maybe this is so in a loose and popular sense, but not in a strict, philosophical sense.

Further, the thought experiment is inconsistent with widely accepted Maximality Principles. The notion of maximality is an important one, and, as explained in Chapter 5, there seem to be two different, but closely related maximality principles relevant to assessing Bailey's thought experiment. We cite two physicalists who embrace one or both principles. The first is embraced by Hud Hudson and the second by Bailey himself.[28]

> *Individual Maximality* (IM): *x* is an individual maximal object just in case, necessarily, *x* is an F and there are no proper parts *p* of *x* that are Fs.
>
> *Property Maximality* (PM): A property P is maximal *iff* necessarily, for any *x* that is F, there is no *y* such that *x* is a proper part of *y* and *y* is F.

IM is Hudson's way of stating the principle, and takes as substitutions for *x* human persons, persons, and living human organisms. For example, IM implies that if *x* is a person, then *x* has no proper parts that are persons. PM is stated precisely as Bailey presents it and his main example of a maximal property is being conscious. As explained in Chapter 5, it follows that

> *C-Maximality*: Being consciousness is maximal *iff* necessarily, for any *x* that is conscious, there is no particular *y*, such that *y* is a proper part of *x*, and *y* is conscious.

It should be clear that all the iterated Bradleys after First-Bradley violate IM (assuming Bradley is a person throughout the transformation process and that angels are persons, then the iterated Bradleys are persons (or the iterated Bradley is a person) and have persons as proper parts). The thought experiment also violates Bailey's own statement of PM which, remember, is a necessary truth. Assuming that Bradley is conscious at every stage of the process, in all stages besides First-Bradley, he is conscious and has proper parts (angels) that are also conscious. By violating IM and PM, it would seem that Bailey's thought experiment is in deep trouble, even by his own ontological commitments.

Finally, briefly put, Bailey's understanding of living organisms, especially as exhibited by the transformation of Bradley, will have serious difficulties accounting for the unity of consciousness (especially for Final-Bradley) and the reality of evidential near-death experiments. We have discussed these elsewhere in Chapter 6 and below and will not repeat the problems here.

---

[28] See Hudson, H. (2001). *Materialist metaphysics of the human person* (p. 121). Ithaca, NY: Cornell University Press. Cf. 26–27, 37–38, 101–105, 117, 121–122, 127–128. Bailey, A. M. (2014). You needn't be simple. *Philosophical Papers*, *43*(2), 145–160, especially 145, 151–156.

Fourth, Bailey defends an implausible view of categories, yet he needs this view to do important work in defending contingent physicalism. As noted above, in developing his ontological categories, Bailey locates himself squarely in the realist camp and claims to be doing old-fashioned categorial ontologizing in the grand tradition of such activity. However, there are two features of his categories that seem to be quite problematic: (1) the contingency of (at least some) items relative to their categories and (2) the identification of important nodes in purely negative terms. Both of these features are important for Bradley's defense of contingent physicalism. Let's examine these in order.

Regarding (1), for various reasons, e.g., to defeat the modal argument for substance dualism by claiming we are, possibly, disembodied, Bailey needs to employ contingent physicalism. He accomplishes this in two ways. First, his presentation of categories isolates their modality—the necessity/contingency of items belonging to their category—and he claims that contingency prevails, at least in the subcategories of Thinking Material Object and Thinking Immaterial Object. Second, his thought experiment about Bradley's transformation is confused. If we understand him correctly, the contingency of the relevant subcategories is the ontological and epistemic grounds for the intelligibility of the thought experiment.

Unfortunately, categorial contingency is an extremely odd view that requires substantial warrant to be accepted. To our knowledge, no realist about categories in the history of philosophy up to the present has even considered, much less advanced such a view.[29] It is widely acknowledged that the great realists about the categories, e.g., Aristotle, Bergmann, Butchvarov, Chisholm, Grossmann, Loux/Crisp, Lowe, and Hoffman/Rosenkranz to name a few, all held that members of categories necessarily belonged to them.[30]

We are not simply collecting votes here. Rather, these and all of the other realist philosophers whom we know agreed on the very nature of what it was to develop a set of categories and what a category was: The goal was to discover a set of exhaustive, (often) mutually exclusive (except for the transcendentals), most basic divisions or reality into which all entities necessarily and essentially fell. A category was constituted by what its members shared essentially in common, and a category often provided essential or necessary and sufficient truth conditions for something to belong to it. Given this near unanimous understanding among categorial realists, as close to unanimity as one finds in philosophy, the idea of separating out a modal question would be otiose. The modality of membership—necessity—was already implicit in the very notion of what it was to be a category in the first place.

---

[29] For a possible exception to this claim, see Westerhoff, J. (2005). *Ontological categories: Their nature and significance*. Oxford: Clarendon Press.

[30] Aristotle. (1963). *Categories* (J. L. Ackrill, Trans. with notes). Oxford: Clarendon Press; *Metaphysics IV*; Bergmann, G. (1967). *Realism: A critique of Brentano and Meinong* (pp. 4–27, 101, 141). Madison, WI: University of Wisconsin Press; Butchvarov, P. (1995). Categories. In J. Kim & E. Sosa (Eds.), *A companion to metaphysics*. Oxford: Blackwell; Chisholm, R. M. (1989). *On metaphysics* (pp. 162–168). Minneapolis, MN: University of Minnesota Press; *A realistic theory of categories*. (1996). Cambridge: Cambridge University Press; Grossmann, R. (1983). *The categorial structure of the world* (pp. 3–18). Bloomington, In: Indiana University Press; *The existence of the world* (pp. 1–3, 46–47, 91). (1992). London: Routledge; Loux, M. J., & Crisp, T. M. (2017). *Metaphysics: A contemporary introduction* (pp. 10–15). New York: Routledge; Lowe, E. J. (2006). *The four-category ontology* (pp. 3–23). Oxford: Clarendon; Hoffman, J., & Rosenkrantz, G. S. (1994). *Substance among other categories* (pp. 14–21). Cambridge: Cambridge University Press; *Substance: Its nature and existence* (pp. 46–50). (1997). London: Routledge. We are not settled on exactly where Peter van Inwagen comes down on this issue. Many things he says indicate he is a realist about the categories and holds that membership is necessary (though some categories are contingent in that they have members in some worlds and are empty in others). On the other hand, while he claims that ontological categories must in some sense be modally robust, he claims that he has yet to spell out precisely what this means. See van Inwagen, P. (2014). *Existence: Essays in ontology*. Cambridge: Cambridge University Press, chapter 9.

Bailey seems to be virtually alone in his approach. This does not automatically render his view of categories unjustified. But surely, if he is going to break with such a homogeneous pack, he owes us at least a brief account of why his approach is justified and why the majority view can be abandoned. But Bailey fails to offer such an account. He simply asserts his categorial contingency thesis. This fact is made even more troubling by something Bailey himself affirms.[31] In discussing the development of categories as an aid in probing the metaphysics of human nature and answering the question of what we are, Bailey rightly distinguishes the categorial approach from other approaches that provide true, but the wrong kind of answers to count as categorial, e.g., we are each no more than 4300 years old, have at least one great-grandparent, and so on. These answers fail because they classify us in virtue of contingent features or features that don't get at the specific, deepest notion of what kind of thing we are. However, it seems that Bailey's contingent categorial membership fails in just this sort of way.

Regarding (2) (using purely negative descriptions of different subcategories), Bailey rejects an argument for substance dualism based on the fact that we have a clear, adequate concept of being a soul or spirit, e.g., when we say that God is a spirit.[32] Briefly, the argument claims that we have a clear concept of a soul/spirit, and for that to be the case, we need a distinct positive concept of such. If thinking matter is metaphysically possible, then the properties that constitute consciousness can characterize either a material object or a spirit. Thus, we cannot appeal to those properties for a distinct positive concept of either. Regarding a conscious material object, we have distinct positive concepts about what makes it physical, but we lack such for a spirit. Characterizing a spirit is either vacuous or by way of negation, and neither provides us with what we actually possess, viz., a distinct positive concept of a spirit. Thus, thinking matter is impossible.

At this point, it may be useful to provide a relevant discussion about the adequacy of a wholly negative characterization of a spirit offered by Joshua Hoffman and Gary S. Rosenkrantz.[33] They define a soul "as a non-spatial substance that has mental properties." Next, they consider this objection that the very concept of a soul is unintelligible:

> A first argument protests that the nature of a soul cannot be fully explained without describing it in *negative terms*, i.e., as *unlocated*. We reply that there are intelligible *physical* entities whose nature cannot be fully explained unless they are described in negative terms. For example, to fully explain the nature of a photon, a photon must be described as having *zero* rest mass (i.e., as *not* having rest mass). Of course, the basic nature of a photon includes certain other, *positive* attributes.

---

[31] Bailey, p. 5.

[32] See Moreland, J. P. (2013). A conceptualist argument for substance dualism. *Religious Studies*, *49*(March), 35–43. Cf. Bailey, pp. 16–22.

[33] Hoffman, J., & Rosenkrantz, G. S. (2002). *The divine attributes* (pp. 41–43). Oxford: Blackwell.

> But the same is true of the basic nature of a soul. It includes such positive attributes as *being a substance* and *being conscious* (or *being capable of consciousness*). So, if it is argued that souls are unintelligible because their basic nature is *wholly* negative, then such an argument is **unsound**.[34]

Two points are important. (1) They seem to assume that if a concept (or entity?) is wholly negative, then it is unintelligible. They go on to say that souls are explained (described? conceptualized?) in positive terms and, thus, are at least partially positive. It is this positive characterization that renders the concept of a soul intelligible. (2) The content of that positive conception captures the basic nature (kindedness) of a soul and includes the categorial feature of being a substance and the essential properties of mentality/consciousness (which are used interchangeably).

Obviously, if successful, this argument would be bad news for Bailey. So he provides defeaters for this sort of argument. Evaluating Bailey's counterargument is not relevant here. But one key defeater he raises is pertinent to our critique of (2). Bailey claims that the Positive Requirement (one needs a distinct positive concept of something to have a clear, adequate concept of it) is false and provides a counterexample to it. He claims that we have an adequate concept of the abstract, but no positive concept of it. Our concept is by way of the *via negative*.

We have two responses to this counterargument. First, it is false. We do, in fact, have a positive concept of being abstract or, more importantly, of abstract things—being abstract is being necessary, being outside space and time, necessarily standing in such and such internal relations to certain entities (e.g., necessarily, redness stands in the internal relation darker than to yellowness). And we can cite distinct positive features of abstract objects, e.g., the number two is essentially even, it is a whole integer that stands mathematically between the numbers one and three, and so on.

Second, the most employed strategy in describing abstractness is by way of listing examples of abstract objects, e.g., various logical relations, properties, numbers. Each member of the list has positive characteristics that are the basis of satisfying the Positive Requirement.

Bailey also claims that we can easily say what, e.g., a material or immaterial object is even if it is characterized by way of the *via negative*. How? According to the Chart Ontologist, one can say what something is by pointing to the relevant node on the chart of categories. Thus, to say what a spirit is requires only that one point to the node Thinking Immaterial Object.

Again, we think this solution fails. While we cannot enter into a discussion of the pros and cons of the *via negative*, we note that in our view, one must know something positive about what an entity is before one can say what it is not. The *via negative* fails in this regard and, without positive knowledge, it leads to skepticism about the nature of the object characterized merely in a negative way.

Moreover, Bailey's strategy seems merely to kick the can down the road. In solving the problem of not being able to say what a spirit is without having a positive concept of a spirit, Bailey claims that by pointing to the node Thinking Immaterial Object, that adequately says what it is to be a spirit. But it seems clear that the node exhibits the same problem of lacking a positive concept of "immaterial" as did the characterization of a spirit as an immaterial or incorporeal spirit that is not material. Unless we are missing something, we don't see a solution here to the initial concern. Until we do, we judge this strategy to be a failure.

---

[34] Hoffman and Rosenkrantz (2002), pp. 42–43. Italics and bold are theirs.

8.2.1.2.3 An Addition Response to Bailey's via Negativa

Before leaving this topic, there is an additional consideration that we believe presents serious difficulties with Bailey's claim that pointing to a purely negative node—Thinking Immaterial Object—is all one needs to do to know what a soul/spirit is. Let us use TMO and TIO for Thinking Material Object and Thinking Immaterial Object, respectively. Adolf Reinach has observed that we often make these four judgments:[35]

(i) *Positive Judgment about a Positive State of Affairs*: It is the case that rose *r* is red.
(ii) *Negative Judgment about a Positive State of Affairs*: It is not the case that rose *r* is red.
(iii) *Positive Judgment about a Negative State of Affairs*: It is the case that rose *r* is non-red.
(iv) *Negative Judgment about a Negative State of Affairs*: It is not the case that rose *r* is non-red.

In all four cases, a conscious side is in some way related to an objective side. Let's focus on (ii) and (iii). According to Reinach, there are two specific preconditions, two position-takings for being justified in asserting (ii).

> *The psychological precondition*: There is an intellectual position-taking about some other state of affairs (we apprehend r's being yellow) that is in conflict with the original state of affairs (r's being red).
>
> *The epistemic fundament*: In approaching the original state of affairs (r's being red) in light of the second state of affairs (r's being yellow) and grasping the conflict, the later appears in a new light, viz., as having an evidential character relative to the former.

On this basis, we are justified in affirming (ii) (It is not the case that r is red.).

In order to avoid quantifying over negative states of affairs as in (iii), Stephen Mumford argues that in all cases of (iii), we reduce them to (ii), placing the negativity on the consciousness side.[36] Applied to Bailey, if we take TIO to be a negative judgment about a positive state of affairs (In some world w with thinking things, it is false that a material thing is thinking in w (or it is false that TMO has these items in w).), then we aren't capturing Bailey's view that pointing to a purely negative node—Thinking Immaterial Object—is all one needs to do to know what a soul/spirit is. As far as we can tell, Bailey's pointing strategy is not a mere expression of negativity, but a positive assertion if it is going to capture "what a spirit/soul is."

So, we need to examine (iii). For the sake of argument, let us grant the existence of negative states of affairs. Regarding positive judgments about negative states of affairs, Reinach claims that such judgments have two preconditions: (1) *a psychological*

---

[35] Reinach, A. (1982). On the theory of the negative judgment. In B. Smith (Ed.), *Parts and moments: Studies in logic and formal ontology* (pp. 315–377). Munchen: Philosophia Verlag, especially 315–54.

[36] Mumford, S. (2007). Negative truth and falsehood. *Proceedings of the Aristotelian Society, 107*, 45–71.

precondition: Focus on the negative state of affairs as such (3 is not smaller than 2) and put it into doubt; (2) *an epistemic fundament (foundation for the positive judgment)*: the apprehension of another, positive state of affairs (3 is larger than 2) such that the two states of affairs are necessarily "bound up" together and the latter is the epistemic foundation for the former.

If (iii) is what Bailey is expressing in his node-pointing strategy, then what is this positive epistemic foundation for the judgment "in w, it is true that there are thinking things that are non-material (or it is true that there are items in TIO)"? He needs a positive state of affairs to be "bound up" with "immaterial." But given his *via negative* approach, this is precisely what Bailey cannot provide.

Bailey could respond that our argument begs the question against his approach. But even if the readers disagree with the details of Reinach's analysis, we believe we learn something from it: *Negative judgments and negative states of affairs are privations or absences, and one cannot understand or have an epistemic fundament for them without possessing an understanding, especially an evidential one, of the associated positivity*.

8.2.2 Clarification and Defense of Premise (3)

8.2.2.1 Conceivability and Possibility

Historically, the main position on how we gain modal knowledge as expressed in (3) (Possibly, I exist and no POs exist.) appeals to an epistemic or truth-tracking connection between something being conceivable (or rationally intuitive) and its being possible (necessary, etc.).[37] Unfortunately, there are at least fourteen different notions of "conceiving," many of which are widely disparate.[38] As a result, even though the notion of "rational intuition" also exhibits different interpretations, nevertheless, it is not as unwieldy as "conceiving" and, further, with clarifications, it is more appropriate for our views in modal epistemology.

Before leaving the topic of conceiving there are two important issues that are usually discussed in association with it. So, while we shall adjust these issues in appropriate ways to express our clarified notion of rational intuition, now is a good place to mention them. First, there is an important distinction between strong (positive) conceiving and weak (negative) conceiving. Limiting our focus to possibility, if s strongly conceives some proposition P to be possibly true or some state of affairs S to be possible, then s "sees" that P is possibly true or S is possible. Strong conceiving confers positive epistemic status on P or S. If s weakly conceives P to be possibly true or S to be possible, then s fails to "see" that P is impossibly true or S is incompossible.[39] In our comments below, we will employ positive rational intuition and different specifications of it.

---

[37] Cf. Evnine, S. (2008). Modal epistemology: Our knowledge of necessity and possibility. *Philosophy Compass*, 3/4, 665–670; Yablo, S. Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, *53*(1), 1–42.

[38] Gendler, T. S., & Hawthorne, J. (Eds.). (2002). *Conceivability and possibility* (pp. 7–8). Oxford: Clarendon; cf. Tidman, P. (1994). Conceivability as a test for possibility. *American Philosophical Quarterly*, *31*(4), 297–308.

[39] Van Cleve, J. (1983). Conceivability and the Cartesian argument for dualism. *Pacific Philosophical Quarterly, 64*, 35–45.

Second, another distinction is between *de re* and *de dicto* conceivings, the former being directed toward an entity, e.g., a substance, and the latter toward a proposition.

8.2.2.2 Modal Epistemology, Rational Intuitions as Seemings and Adequate Intuitive Presentation

In order to clarify our own concept of what confers positive epistemic status on modal claims whether they exhibit the status of being *prima facie* justified, *secunda facie* justified, or (usually) defeasible modal knowledge, it seems best to survey three representative samples of modal epistemology to give a proper context for our own views.[40] And such a survey can provide different approaches to countering modal skepticism, e.g., by giving different explanations for modal error.

Consider George Bealer's statement: "Intuition is the source of all *a priori* knowledge."[41] Setting aside purported *a posterior* modal knowledge, we take Bealer to be on the right track. Assuming this as a launching point, we, first, note Chisholm's comments that the concepts expressed in a class of important words are fundamentally used in a primitive phenomenological way: "seems," "appears," "looks."[42] Accordingly, we take phenomenological employments of "seemings" or "appearings" to express "intuitions." Stated adverbially, a rational seeming is a way of being appeared to by a purported intentional object.

In our view, any account of modal knowledge (evidence, justified belief) must answer two questions:

> *Q1*: What is the character of adequate intuitive acts, the intentional objects of such acts, and the tie that connects them in cases of modal knowledge?
>
> *Q2*: How does one account for modal error?

As we will see, some reject *Q1* altogether and replace "intuitive acts" and "intentional objects" with causal or reliabilist considerations while others retain (at least) some notion of intuition as a necessary but not sufficient condition for modal knowledge to be supplemented by externalist factors, e.g., modal reliabilism. With this in mind, let us examine three versions of modal knowledge.

8.2.2.2.1 *Timothy O'Connor*

According to O'Connor, modal beliefs are woven into our concept of the world. From our theorizing, e.g., advancing scientific theories, our cognitive activities are replete with modal concepts. Thus, he adopts *Robust Modal Realism*: The world contains primitive modal features.[43]

---

[40] Elsewhere, one of us has assessed Bonjour's modal epistemology. See Moreland, J. P. (2001). *Universals* (pp. 123–126). Montreal and Kingston: McGill-Queen's.

[41] Bealer, G. Modal epistemology and the rationalist renaissance. In Gendler & Hawthorne (pp. 71–125), especially 73.

[42] Chisholm, R. M. *Theory of knowledge*, pp. 26–34, 39–42. So understood, it follows that there is a unique what-it-is-like to conscious states taken as ways of being appeared to. Often, philosophers limit such phenomenal textures to sensations, but we take them to constitute the content side of propositional attitudes such at thinking that P. For more on this, see Pitt, D. (2004). The phenomenology of cognition or *What Is It Like to Think That P? Philosophy and Phenomenological Research*, *LXIX*(July), 1–36.

[43] O'Connor, T. (2012). *Theism and ultimate explanation* (pp. 1–62). Malden, MA: Wiley-Blackwell.

These claims raise an epistemological problem, especially in light of modal error: How do we come to have modal knowledge? O'Connor's answer employs a version of truth-maker theory according to which for every truth—modal or otherwise—there is some entity that is its grounds and whose obtaining metaphysically necessitates its truth.

Two initial observations are in order. First, it seems that O'Connor's sole focus in modal epistemology is the *de dicto* knowledge of a modal proposition's truth value (though he does believe in modal facts). He regularly talks about modal beliefs, concepts, judgments, or theoretical explanations that tacitly assume modalities that are not transparent to an idealized grasp of the relevant concepts. Here, it is concepts/propositional content and not mind-independent modal entities that are known.

Second, O'Connor rejects the role of rational intuition in modal knowledge when understood as a form of seeming, directly perceiving, and so forth. He explicitly equates intuition with belief.[44] He also rejects such seemings on empirical grounds, and he faults Bonjour's account of modal knowledge for employing rational "seeing" which, says O'Connor, is not explicated beyond a mere disposition to believe.[45] Further, O'Connor faults earlier rationalist employment of intuition as a metaphor for "directly seeing the necessity" since it is too strong. Rather, once we embrace modal fallibilism, we should replace this metaphor with a strong disposition to believe certain propositions to be necessary, a disposition often fueled by a hidden, underlying commitment to the explanatory role of such necessities.

As we will see shortly, we also think O'Connor's exclusive focus on *de dicto* modality results from his reductive, causal chain analysis of acquaintance or direct awareness that, at least in the case of indexicals and demonstratives, seems to require eschewal of *de re* reference on which such reference is grounded. While we do not agree with much of John Campbell's account of consciousness and reference, we think he got it right when he said, "Knowledge of truths depends on acquaintance with objects. For it is acquaintance with objects or properties that provides our knowledge of reference."[46] Campbell goes on to argue that employment of causal chains in this context is best explained in terms of propositional knowledge and, thus, a loss of intentional direct awareness as a fundamental source of knowledge. For Campbell, this makes it unclear as to how or why consciousness is relevant to perceptual knowledge.

One could respond that (at least for) demonstrative knowledge, the causal chain analysis does, in fact, have a fundamental ground for propositional demonstrative knowledge, viz., an "appropriate" causal chain between the object and some neural or other scientifically characterized nonconscious mechanisms. But this response gives no role to consciousness in providing direct awareness of an intentional object and fixing reference to it. This is consistent with the reduction/elimination of consciousness altogether or rendering it epiphenomenal, at least with respect to our perceptual knowledge of objects.

Similarly, one could hold that the previous analysis is wrong because it selected the wrong relata for the causal chain. Accordingly, the causal chain between object and *the intrinsic aspect of the experience of the object* is what makes the latter of or about the object. But as Campbell observes, now it is hard to see why it is the experience that matters: "[I]f a causal link from the object to the experience is what is important, why would it not be good enough to have a causal link from the object to some non-conscious aspect of the perception?"[47]

---

[44] O'Connor (2012), p. 149, note 11.

[45] Ibid., pp. 9, 10.

[46] Campbell, Consciousness and reference, pp. 648–662. The quote is on page 650.

Turning to O'Connor's positive account of modal knowledge, four features are important for our purposes: the role of causality, reflective equilibrium and refinement, the tie between modal facts and modal knowledge, and sources of modal error.

First, O'Connor adopts a causal component for knowledge—including modal knowledge—as a necessary condition.

> *Causal Modal Knowledge*: Modal knowledge entails some sort of appropriate causal connection to modal facts.[48]

Second, in addition to a causal component, reflective equilibrium and refinement are necessary components of modal knowledge.[49] According to O'Connor, we begin our quest for developing a theory of modal knowledge by starting with the insight that modal beliefs play a unique, ineliminable role in common sense and in a theoretical understanding of things. We begin with certain core "data" and through reflective equilibrium and refinement, we extend outward to a fuller theory of modal knowledge.

For example, we have a strong propensity to accept certain obvious modal propositions, and we recognize that our explanatory schemes and theoretical rationality are built upon taking such basic modal notions as fundamentally legitimate. Moreover, we recognize that our perception of the actual is imbued with an implicit grasp and acceptance of the modal character of reality. As we build from this "ordinary folk" perspective, we develop theories of modal knowledge that extend the range of our modal beliefs and allow for correction, refinement, and extension of our modal judgments.

Regarding the tie between modal facts and knowledge, O'Connor appropriates truthmaker theory:[50] Modal facts ground modal truths and by way of a causal chain, trigger a "strong propensity" to accept certain modal truths, a "disposition to strongly believe" them, particularly those involving necessity and impossibility. To give an account of why such dispositions are generally truth-tracking, we need to reflect on the causal origin of our fundamental modal convictions. Any account cannot be causally independent of the relevant truthmakers.[51] To fill out a sketch with more details, O'Connor presents a just-so evolutionary story:

> An evolutionary advantage accrued to cognizer-types that readily assent to the *actual* truth of core logical and mathematical principles and that systematize the world in terms of natural kinds; some such cognizers in our ancestral history were selected in part owing to this fact; and *the truthmakers for these actual truths are none other than their modalized counterparts*.[52]

---

[47] Campbell, p. 652.

[48] O'Connor, pp. 8–10, 50, 51, 58–59.

[49] O'Connor, pp. 41, 42, 43, 49.

[50] O'Connor, pp. 8, 10.

[51] O'Connor, pp. 58–59.

[52] O'Connor, p. 59.

Perhaps not exhaustive, but the following are the major sources of modal error for O'Connor:

- Deviant causal chains from modal fact/truth to modal propositional dispositions/attitudes.
- Unreliably produced, non-truth-tracking modal dispositions.
- Disagreement regarding the modal implications of non-modal facts. (Could the embodied person Socrates have been an alligator?)
- General theoretical disagreements where the role of imagination is far from clear. (Is time travel possible?)
- Extending modal cognitive processes to areas we cannot imagine in sufficient detail or which are such that they involve controversial philosophical possibility claims, e.g., remote and complicated unactualized possibilities.

This précis does not do justice to the full extent of O'Connor's reflections, but this will do for our purposes. Our goal in this section is to sketch out main and diverse versions of modal epistemology in order to show the resources available to one who embraces premise (3). Many accept O'Connor's approach and, accordingly, could use it with adjustments to defend (3). Still, for three reasons, we are unpersuaded by his view.

First, O'Connor mistakenly makes modal knowledge about *de dicto* modality with no room for *de re* modal knowledge as the epistemic grounds for the former. We recognize that O'Connor advances a version of epistemic access-externalism, so he may not be bothered by this criticism. But we think his doing so is uninformed. Surely, commonsense "folk epistemology" is internalist in some way or another, and we think one should stick to that unless there are substantial defeaters that override its *prima facie* justification. Just as seeming to see red while looking at an apple provides good justification for believing the apple is red, so seeming to be aware of modal entities does the same thing.

But O'Connor's externalist account of modal knowledge is revisionist with respect to common sense. To cite one example, when he claims that our perception of the actual is imbued with an implicit grasp and acceptance of the modal character of reality, he will have to unpack "grasp" in terms of causal chains and unconscious physical dispositions. If this is correct, it is hard to see how there could be a perceptual experience of the actual since a perceptual sensation would seem to be an event at the subject-end of the causal chain standing between that subject and the appropriate actual entity. Not only does such a chain seem to cut off the perception from the actual world "out there," but it also seems to make the intrinsic intentionality of the perception superfluous for "hooking it up" with its intentional object.

The only two defeaters for rational intuition we can find in O'Connor's writings are these: (1) Philosophers who appeal to rational intuition fail to explain, e.g., how "seeing" redness and greenness allow us to "see" their mutual exclusivity. This is not explicated beyond claiming a mere disposition to believe the exclusion. (2) Those who appeal to rational intuition, at least historically, adopt infallibilism for modal knowledge and fail to account for (purported) modal error.

In response, (2) is simply not applicable in the current setting. Even Edmund Husserl acknowledged the defeasibility of modal awareness and propositional attitudes epistemically grounded in them, and he proffered a clear account of modal error. Regarding (1), Husserl provided the most detailed answer to this problem in the history of philosophy. We will examine his views below.

Second, due to his inadequate employment of irreducible intentionality, O'Connor turns to a causal account as an essential aspect of his overall view of modal knowledge. Unfortunately, (efficient) causality and appropriate causal chains are irrelevant to sensory (e.g., visual) and modal evidence, justified belief and knowledge.

The obtaining of such a causal relation/chain between an entity known—especially one known by a direct intuitive awareness—and the knowing subject is not necessary for knowledge. Two considerations show this to be the case. (1) Consider a world in which relevant entities have causal powers and such chains obtain. Now suppose that God decides to shift this world to an occasionalist one every Sunday and is the sole direct cause of everything. If one was appeared to redly while looking at a red apple's surface (or used the concept of a dog to think about one) from Monday through Saturday, then if one had that same sensory state (or concept) on Sunday, one would still be seeing the apple's surface or thinking about the dog *sans* the relevant causal chain. If someone says that God would still be doing the relevant causal work, this won't do. What exactly does God cause in the subject such that it "hooks up" with the right object? If God were to cause a sensation of a gray elephant while one is looking at the red apple, the sensation simply would not be about the apple. It would be about the elephant.

(2) Another way to see the unnecessariness of a causal relation is to consider a case presented by Hillary Putnam.[53] In order to get rid of irreducible intentionality and replace it with wide mental contents and a causal chain analysis, he invites us to consider a being on some planet with no trees who suddenly gets a mental image of what we would describe as one of a tree. Putnam claims that the image is not a representation of a tree because mental states, e.g., images, are not intrinsically representational and there is no (appropriate) causal chain from a nonexistent object to the image.

In responding to Putnam's thought experiment, Geoffrey Madell grants the lack of intrinsic intentionality for the mental image but goes on to give the reason why.[54] It is not due to the absence of a causal connection; rather, it is because no sense image is intrinsically referential. Images become representational when the subject stipulates or thinks of it as representing something, in this case, a tree.

In our view, this whole discussion is deeply confused. Due to a sensation's or concept's intrinsic character (available to second-order introspection), it is of what it is of. It functions as a natural sign. You can't have a tree sensation without it being of a tree (even if no tree exists or caused it.) Moreover, (at least clear) intentional mental states are necessarily of their intentional objects. In the alien's case, his image is of a tree, but unfortunately, since no trees exist, the image fails to "hit" a real target.

Regarding Madell's response, he seems to liken visual images (we cash them out adverbially) to words and this is wrong. The word–world connection is arbitrary and conventional. The visual sensation or conceptual relation to an intentional object in the world is necessary. As we will see in our examination of Husserl, (especially) certain kinds of sensations and propositional attitudes are intrinsically intentional and necessarily of their intentional objects, and this is why we know what we are thinking about or taking ourselves to see by a simple act of focused attention on the intentional state itself. This ability is hard to explain if a causal chain or stipulative analysis is embraced. Either approach entails that there is nothing intrinsic to the mental state of which one could be aware and that makes the state of its intentional object.

---

[53] Putnam, H. (1981). *Reason, truth and history*. Cambridge: Cambridge University Press, Ch. 1, especially 4.

[54] Madell, *The identity of the self*, p. 17.

In addition, the obtaining of such a causal relation/chain between an entity known and the knowing subject is not sufficient for knowledge. As O'Connor acknowledges, causal chains can always be deviant. This is why causal-connection advocates require that the chain be "appropriate." It is notoriously difficult to cash this out in a noncircular way. But waving this problem, the dialectical process of refining and giving content to "appropriate" inevitably involves appealing to the intrinsic character and intentionality of the relevant mental state to guide the deliberative process. An analysis of "appropriate" is advanced, and counterexamples offered. A new, patched-up analysis is offered and new counterexamples are provided, and so on until an allegedly acceptable final analysis is reached.

What is it that guides the process, providing the intuitive insight that a given analysis adequately solves the relevant counterexamples and providing the insight for generating new counterexamples? It is our (at least implicit) understanding of the intrinsic intentionality of the relevant mental state, of what it is supposed to be about. If we are right about this, then the relevant intrinsic intentionality is irreducible/ineliminable, and it is conceptually and epistemically basic while the causal-chain analysis is derivative.

Besides the fact that a causal chain analysis is neither necessary nor sufficient for direct intuitive knowledge, as we saw above in our mention of John Campbell, it is entirely unclear as to why or how consciousness itself (and intentionality) is involved in causal chain accounts. This seems to be a real problem with the analysis.

Finally, advocates of "intuition as direct awareness" view face, and have adequately responded to, the problem of modal fallibility. But causal chain analyses face an associated difficulty lurking in the neighborhood: the problem of incorrigible or infallible modal judgments, along with the directness/immediacy of modal seemings. Granting widespread modal defeasibility, there still seem to be cases of modal infallibility, e.g., the transitivity of identity and certain other relations like the law of noncontradiction (with apologies to advocates of deviant logics). If causal chains are always subject to deviancy, how is one to account for these cases? And what are we to make of the apparent immediacy and directness of our intentional awarenesses? These are difficult questions and causal-chain advocates need to do a better job of answering them.

In sum, we agree with William Lycan who flatly states that "knowledge does not require that one's belief has been caused by the fact known."[55]

8.2.2.2.2 George Bealer[56]

As David Kasmier observes, because Bealer accepts the widely adopted JTB (plus Gettier modifications) analysis of knowledge, knowledge strictly entails "evidence," so a theory of evidence is all that is necessary for a theory of *a priori* knowledge, and a theory of the latter must answer two questions: (1) What are intuitions? and (2) Why are they evidence?[57]

---

[55] Lycan, W. (2018). Redressing substance dualism. In J. Loose, A. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (p. 25). Oxford: Wiley-Blackwell.

[56] Key primary sources are Bealer, G. Modal epistemology and the rationalist renaissance. In *Conceivability and possibility* (pp. 71–125); Philosophical limits of scientific essentialism. *Philosophical Perspectives, 1,* 289–365; On the possibility of philosophical knowledge. (1996). In J. E. Tomberlin (Ed.), *Philosophical perspectives 10: Metaphysics* (pp. 1–34). Cambridge, MA: Blackwell. Key secondary sources are Kasmier, D. (2003). *Husserl's theory of a priori knowledge: A response to the failure of contemporary rationalism*. PhD Dissertation, University of Southern California, pp. 38–50, 96–103; Evnine, S. (2008). Modal epistemology: Our knowledge of necessity and possibility. *Philosophy Compass, ¾*, 664–684; Hopp, W. (2020). *Phenomenology: A contemporary introduction.* New York, NY: Routledge; Piazzo, T. (2007). *A priori knowledge: Towards a phenomenological* explanation (pp. 137–181). Frankfort: Ontos Verlag; Yablo, S. Is conceivability a guide to possibility?

In answering these questions, *de re* modality seems not to play a role; rather, Bealer's theory is cashed out with respect to *de dicto* modal knowledge by focusing on analysis and knowledge of concepts and the truth of modal propositions. His is a theory of evidence as intuitions supplemented by modal reliabilism to explain why and under what conditions certain intuitions are defeasibly truth-tracking.

Intuitions are a basic source of evidence and as such, are primitive and *sui generis*. Rational intuitions are seemings whose objects are modal concepts, modal propositions, and their truth values, and seemings amount to intuitive understandings. Such intuitions do not involve some magical power or inner voice of some mysterious faculty. For example, upon reflection on a De Morgan law, a modal seeming occurs and, upon further reflection, an alethic seeming obtains.

Why are intuitions evidential? According to Bealer, certain intuitions have an appropriately strong modal (i.e., necessitating) tie to modal truth elucidated by way of his version of modal reliabilism. And what is this version? To begin with, the appropriate tie requires properly understanding the concepts involved in the content of our modal intuitions. To possess a concept is to understand it to some degree or other. However, only a certain kind of concept possession—determinate concept possession—entails the needed tie. A subject possesses a concept determinately *iff*

1. The subject at least nominally possesses the concept, and this happens *iff* the subject has a natural propositional attitude (e.g., a belief) toward the proposition which has that concept as a conceptual content.
2. The subject does not do this with misunderstanding or incomplete understanding or just by virtue of attributions practices or in any other such manner.[58]

Condition (2) is called robust or determinate concept possession. When this occurs, one's intuitions are truth-tracking. Determinate understanding of this sort is that mode of understanding that constitutes the categorical base of the metaphysical possibility of truth-tracking intuitions.

Bealer acknowledges that historically speaking, rationalists accepted the following equivalence.

E: For all propositions P, necessarily, P is *a priori iff* P is necessary.

But since Kripke, scientific essentialist considerations have presented a challenge to *E*. Kripke famously showed the equivalence in *E* fails in both directions. Modal error occurs regarding necessary *a posteriori* truths (e.g., necessarily, water is H_2O) with respect to an illusion of possibility in cases where those in an epistemic duplicate community (e.g., Twin Earth) say something true (their aqueous stuff is *xyz*), but this doesn't imply that water itself is not H_2O. Similarly, modal error occurs regarding contingent *a priori* truths with respect to an illusion of necessity when one confuses a description (e.g., "stick S at time t_0") as a definite description vs. as a reference fixer.

---

[57] Kasmier, p. 38.

[58] Bealer, Modal epistemology and the rationalist renaissance, p. 102.

Bealer's way out of this problem is to say that *E* still holds for a certain class of propositions, viz., those that are semantically stable. According to Bealer, there are three relevant classes of concepts at play: (1) categorial concepts (identity, property, relation, stuff, compositional stuff, functional stuff); (2) content concepts (phenomenal concepts, concepts of psychological attitudes); (3) naturalistic concepts (natural kinds involved in scientific essentialism such as water=H_2O). The concepts in (1) and (2) are semantically stable and those in (3) are not. A proposition is semantically stable *iff* its meaning remains invariant across epistemic duplicate communities.

Since the concepts employed in the traditional *a priori* disciplines (logic, mathematics, first philosophy) and those that express the relevant mental concepts are semantically stable, *a posteriori* necessities do not arise in these areas. Thus, Kripke's arguments apply to (3) and have no significance to (1) and (2). Below, we will briefly mention two current strategies for responding to Kripkean problems with *E*, but for now, we want to keep our focus on Bealer's modal epistemology.

In closing out our précis of Bealer's thought, we list his take on the major sources of modal errors along with a brief comment regarding his assessment of the modal argument for SD:[59]

- Confusing epistemic and metaphysical possibility.
- In the context of scientific essentialism, Stephen Yablo's class a belief-based errors: mistaken *a posteriori* beliefs (e.g., that Hesperus≠Phosphorus).
- Stephen Yablo's class b beliefs: mistaken beliefs about the relationship between *a posteriori* beliefs and their associated modal truths (e.g., denying that if Hesperus=Phosphorus, Hesperus could not outlive Phosphorus).
- Those resulting from local categorial misunderstanding according to which categorial mastery of a concept is interrupted due to being locally disrupted. This occurs when one has the right primary categorial concept but does not know it applies to a local range of cases, perhaps due to having the wrong sub-categorial concept (e.g., having the categorial concept that water is macroscopic stuff but selecting the subcategory "functional stuff" when applying it to water in a local situation).
- Those resulting from categorial misunderstanding (e.g., taking water to be a functional concept rather than a compositional-stuff concept).

Finally, for two reasons, Bealer rejects the modal argument for SD.[60] First, regarding intuitions that support the claim that one's disembodiment is possible, Bealer retorts that many fail to have such intuitions and, in fact, many report having the opposite intuitions, so these cancel out. Second, since "I" is a paradigm case of semantic instability, the question is raised as to what subcategory of thing is hit by "I." Granting that "a person" is the primary category, we just don't know if the correct subcategory is "essentially embodied person" or "disembodiable person."

Bealer has presented a careful, sophisticated account of modal evidence (justified belief, knowledge) and modal error. Given the current widespread acceptance of different versions of epistemic externalism, we think his views should be appealing to many philosophers, and surely, it shows that a plausible account of modal evidence, et al., and error is available. Still, we think we can do better, and we have four problems with Bealer's account.

---

[59] Due to space considerations, we omit Bealer's employment of *a priori* stability. See Bealer, Modal epistemology and the rationalist renaissance, pp. 104–105.

[60] Ibid., pp. 114–116.

First, it is entirely *de dicto* in that the objects of modal knowledge are concepts, propositions, and truth-values. This is especially odd, given the importance of rational intuitions as seemings. It is most natural to take such as direct awareness or apprehension of a relevant entity. Given this, we see no reason to limit those entities in the way he does; rather, it seems more fitting to hold that such seemings are also directed at relevant modal entities themselves (e.g., properties, relations) epistemically prior to *de dicto* seemings such that *de re* seemings provide the epistemic grounds for *de dicto* seemings.

Bealer's omission of such *de re* seemings is responsible for his taking what we believe to be a wrong turn in cashing out why intuitions as seemings are evidence appropriately tied to truth. Given the importance of intuition/rational seeming for Bealer, one would naturally expect an elucidation of the appropriate tie in terms of a careful analysis of the nature of seemings, truth, considering, reflecting, propositions, and how they are made available to and are in consciousness, modal entities and the like. One would also expect an account of how these are brought together in synchronic and diachronic mental acts so as to deliver modal knowledge. Instead, Bealer jumps to an externalist account in terms of his version of modal reliabilism. Given his starting point in seemings, it would have been simpler and more natural for Bealer to have continued along an access-internalist path, rather than changing horses in midstream.

Second, the difficulties just surfaced actually make intuitive seemings superfluous to his overall account. To see this, consider functionalist accounts of mental types and tokens. Mental types are functional roles and mental tokens are realizers of those roles. Interestingly, the intrinsic nature of the realizer is irrelevant to making it a token of a particular type of mental event. Rather, the relevant factor is the realizer playing the correct role in an input/output system, whether causal or in a machine-functionalist way. Thus, if the realizer is one with the intrinsic phenomenal texture of being hurtful, that is irrelevant to its being a pain. What makes it a pain is its playing the right role in the functional type being painful.

Applied to Bealer, he takes intuitions to be psychological conditions that serve as indicator evidence for various necessities. It is the reliable processes and correlations that make intuitions epistemically important. Thus, these processes/correlations need not be known or thought about for them to be relevant. All that must take place is that the appropriate reliable connections obtain between the processes that produce modal beliefs and their truth value. Thus, the intuition itself is no *per se* indicator at all. Intuitions are irrelevant *as intuitions*.

Further, given the importance of determinate concept possession, especially in the sense that it entails *understanding* that concept in which proper *categorial mastery* is central, Bealer runs into a problem. Bealer's account depends on what it seeks to establish. Understanding the relevant concept and categorial mastery of it require *a priori* knowledge. That intuitions become evidence, therefore, depends on *a priori* knowledge, so the former cannot constitute the latter.

Third, according to Bealer, intuitions are a basic source of evidence and, as such, are primitive and *sui generis* propositional attitudes. He is quick to distance his view from those who hold that intuitions are a magical power or inner voice of a mysterious faculty. But these claims represent two significant confusions.

(i) The idea of a faculty has a long and steady history and is still employed today.[61] Many contemporary philosophers are irreducible realists regarding notions such as various kinds of powers, dispositions, potentialities, capacities, abilities, etc. Further, historically speaking, the sub-categorial classification of these notions was grounded in the kind of properties they actualized or activities they manifested.[62] Finally, a faculty is a natural grouping of relevantly resembling dispositions. There is nothing particularly odd, spooky, or mysterious about the existence of faculties. Now a faculty of rational intuition (e.g., a natural grouping of dispositions for rational seeming/awareness) may be mysterious within a naturalist, e.g., ontology and anthropology, but Bealer's entire account is mysterious in this sense as are the existence of modal knowledge and the powers required to have it.

(ii) These claims (ironically) represent a category fallacy between a faculty containing dispositions for rational intuitive seemings and the actualized intuitive seemings themselves. It is wrong to conflate and, therefore, identify intuitions and magical powers just as it is to identify a disposition and its manifestation.

Finally, while our account of modal knowledge is primarily directed at defending premise (3) of our modal argument for SD, nevertheless, we want to respond briefly to Bealer's rejection of that premise. Recall that Bealer rejects (3) on two grounds: (a) Many do not have the intuitions expressed in (3) or do have contrary intuitions. (b) "I" is semantically unstable and this obfuscates the subcategory of entity it targets (disembodiable or essentially embodied person).

We will provide important defeaters for (a) latter, so we set it aside for now. Regarding (b), it faces several problems.

(i) As Trenton Merricks points out, the indexical "I" could be removed in (3) and replaced with a premise with no indexicals, e.g., "Possibly, S exists and no POs exist."[63] Acknowledging this point does not require abandoning the metaphysical significance of the irreducibility/uneliminability of "I" or accepting the equivalence in content of (3) and its non-indexical replacement.

(ii) It may be the case that "I" has no conceptual/semantic meaning to begin with. It seems to be the paradigm case of a rigid designator that is purely and directly referential.[64] Bealer could extend his view to include referentially stable and unstable terms, but to our knowledge, he has not done so; consequently, we do not know how he would respond to this objection.

(iii) It may be that "I" has meaning, but in a relevant sense it is semantically stable. Recall that for Bealer, a sentence (proposition, concept) is semantic unstable *iff* the external environment contributes to its meaning. Scientific essentialist concepts are examples of semantic instability. But recall that phenomenal concepts are

---

[61] Perler, D. (2015). *The faculties: A history*. Oxford: Oxford University Press. For an account of the faculty psychology in the 1800s and the factors that lead to its demise, see Reed, E. S. (1997). *From soul to mind* (pp. 22–108). New Haven, CT: Yale University Press.

[62] Ibid., pp. 11–12.

[63] Merricks, p. 84, note 1.

[64] See note 12 above. It does not follow that mental states involving indexical types or tokens do not have nonconceptual intentional contents.

> semantically stable, presumably, because their meanings are in some way fully fixed by something "internal" to the subject. This may mean "metaphysically intrinsic to" or, more likely, something to which the subject has private, direct access.

Given this, here is a plausible account of indexicality—we'll focus on "I"—that may be helpful. In a famous paper, David Kaplan analyzed indexicals such as "I" as follows.[65] Kaplan distinguished two levels of meaning of indexicals: content and character.

Content is the contribution of an indexical term to what is said in a context. The content of "I" varies across contexts of utterance and may be understood as part of a singular proposition composed of the specific referent and the property attributed to it. Character is a rule for using indexical statements and, as such, it may be understood as a function that assigns to any appropriate context of utterance tokening an indexical proposition, the content that would be expressed in that context. As such, it remains invariant across contexts. Thus, character would be semantically stable in Bealer's sense, and while content varies across different contexts, there is nothing in the *external* environment that contributes to its meaning. Therefore, it is disanalogous to scientific essentialist terms in that the latter but not the former exhibit what Bealer takes to be troublesome.

We have examined the accounts of modal knowledge proffered by O'Connor and Bealer. It is time to turn to the account we favor: Edmund Husserl's theory of modal knowledge.

8.2.2.2.3 Edmund Husserl[66]

As David Kasmier observes, typical rationalist accounts of modal knowledge and *a priori* intuition suffer from a failure to provide an adequate phenomenology of rational intuition.[67] More generally, they lack an adequate ontology and phenomenology of modal knowledge by failing to answer two questions: (1) What synchronic and diachronic elements go into the make-up of a piece of immediate *a priori* modal knowledge? What is the essential character of various intuitive acts? (2) How are these elements appropriately connected (synchronically and diachronically) to achieve *a priori* modal knowledge?

Husserl's account answers these questions, along with a careful description of the objects of modal knowledge (dependencies/independencies or compatibility/incompatibility-relations binding objects and their "parts" into complex wholes whose type of unity is grounded in the essences of those object and parts) and the tie between the relevant mental act and the modal object (an object, whether existent or nonexistent, is anything upon which an intentional act may be directed[68]). For every existent object there corresponds the possibility of a set of mental acts and act-qualities (the mode of attending to an object, roughly, the propositional attitude itself) in virtue of which the object in itself comes to be directly present to consciousness and known.

---

[65] Kaplan, D. (1970). On the logic of demonstratives. *Journal of Philosophical Logic*, *8*, 81–98. Harry Cleeveley develops a 2D semantics approach to indexicals and demonstratives that is close to Kaplan's, but he replaces character and content with the primary and secondary intension, respectively. While Cleeveley's approach is a bit reductive, he retains the secondary intensions of indexicals and demonstratives as pure rigid designators. See Cleeveley, H. (2019). *Modal rationalism and the metaphysics of mind*. PhD dissertation, pp. 57–65. London: King's College, especially 60–61.

[66] It is hard to identify a particular location where Husserl treats rational *a priori* intuition and modal knowledge. For those not familiar with his thought, two good introductory books are Russell, M. (2006). *Husserl: A guide for the perplexed*. London: Continuum; Velarde-Mayol, V. (2000). *On Husserl*. Belmont, CA: Wadsworth.

[67] Kasmier, v. We have generally followed Kasmier's order of presentation in chapter 4 because of its clarity.

Given this overview, we turn to a brief description of the epistemological side of modal knowledge followed by the ontological side. For Husserl, the epistemic task is to identify and describe those mental acts/act-qualities through which one becomes fully presented with the modal entities deemed *a priori*. Four items are central to Husserl's epistemology of modal knowledge.

Knowledge and Intentionality. For Husserl, *de re* knowledge obtains when a certain type of immediate intuition/presentation of (direct acquaintance with) the thing itself takes place. Indirect *de re* knowledge involves inference(s) that trace back to immediate presentations. *De dicto* knowledge is the conscious possession of truth/adequation on an appropriate immediate or mediate basis in nonconceptual experience. Immediately based knowledge is the possession of truth in virtue of direct encounter with the facts/things themselves. Mediately based knowledge involves inferences from immediately based knowledge. In contemporary terms, Husserl would agree that *de dicto* knowledge is true warranted belief, but he would cash out "warranted" in this way: It is the possession of immediate or mediate evidence where evidence is provided by a certain kind of presentation in which the thing itself is directly given to consciousness in an act of fullfilment. For Husserl, direct/immediate knowledge is defeasible.

Intentionality is a monadic property (not a relation) that is exemplified by most, if not all, mental states, and it consists in a primitive ofness, aboutness, directedness toward an object or possible object.

Fulfillment and Fulfillment Structures. For Husserl, knowledge involves the recognition of an identity between an object adequately intuited and the object as merely thought. The conscious awareness of this identity is called "fulfillment." More precisely, three conditions must be satisfied in order for epistemic fulfillment to obtain:[69]

> *Intuitive condition*: The intentional object A must be directly perceived or otherwise intuitively presented.
>
> *Conceptual condition*: The intentional object A must be conceived of, judged about, or otherwise conceptually represented in what is usually a relatively empty intention.
>
> *Synthesis condition*: The intuitive and conceptual acts of perceiving or intuitively presenting A must enter into an appropriate synthesis.

Fulfillment is knowledge. And fulfillment is usually the final mental act of a diachronic series of relevantly ordered mental acts that lead a subject from an empty intention to an adequately filled intention in which the object known is directly, intuitively present as it really is. Such a direct awareness of the object provides *de re* knowledge of it and, on that basis, an awareness of the *de dicto* truth value (and as we will see, modal status) of the corresponding proposition. Such a series is called a "fulfillment structure."

---

[68] For a contemporary presentation of Husserlian analysis of intentionality and fulfillment, see Willard, D. (1982). Wholes, parts and the objectivity of knowledge. In B. Smith (Ed.), *Parts and moments: Studies in logic and formal ontology* (pp. 379–400). Munich: Philosophia Verlag; How concepts relate the mind to its objects: The god's eye view vindicated (1999). *Philosophia Christi*, *2*, 5–20.

[69] See Hopp, *Perception and knowledge*, p. 192.

An example may help. Suppose one's wife says there is a shirt in the closet ready for one to wear. Immediately, one forms an empty intention, a vague, general concept of some shirt or other with very few details. But the empty intention has enough content to indicate a series of steps to take to verify or fulfill it. One knows to go to one's closet instead of the garage and to look in the general area where one's shirts and not one's pants hang. So, one embarks on a trip to the closet. One looks where the shirts are and, eventually, one sees a specific short-sleeve, blue shirt with a specific logo on the top left of the front side. One now sees the shirt itself and one has a fulfilled intention at which the shirt itself is directly present to one.[70] One has conceptually independent *de re* knowledge by acquaintance with the shirt, and on that basis, if one compares the shirt as merely thought of with the shirt itself, one can become aware of the truth relation itself and form the relevant proposition expressing *de dicto* knowledge.

Eidetic and Categorial Intuition. Obviously, Husserl did not limit experience and fulfillment within the bounds of sensations. Specifically, he held to two other forms of fulfillment with fulfillment structures appropriate to the nature of the intentional object itself: direct awareness of universals and relations (eidetic intuition or *wessenshau*) and categorial intuition. Each type of object has its own particular mode of being known, of finding fulfillment.

Elsewhere, one of us has given a rigorous analysis of eidetic intuition and we cannot go into that here.[71] Briefly put, Husserl held to an Aristotelian constituent ontology according to which a universal is non-spatially in the concrete particular and, in a different way, the property-instance of that universal (a moment for Husserl), constituting its essence. When one has a direct sensory awareness of a property-instance (a moment) of red, one sees a particular. But by virtue of a sort of gestalt switch, one can attend to the universal itself that is in the instance. And similarly with relations. Categorial intuition is of the objects of formal ontology, e.g., ultimate and subcategories such as object, material object, part, essence, types of states of affairs.

In sum, Husserl believed that one could see a universal with one's very eyes if, of course, a visual quality is in view.[72] But this act of seeing is an eidetic intuition with its own act-character. There are five aspects of eidetic intuition and its relationship to ordinary sensuous perception (construed in a direct realist way) of a moment or concrete particular:

(i) The universal itself is "in"—though not spatially in—the moment.[73]
(ii) We directly perceive the universal via eidetic intuition. Eidetic intuition is precisely parallel to ordinary perception in that through an eidetic intuition the universal is directly perceived. It is placed immediately before the subject.[74]

---

[70] For Husserl, a characteristic of material objects is that one can never have a full and complete awareness of them. There will always be a part, additional angle of awareness or unexamined aspect to them. By contrast, a mental content such as an image is fully, completely present to the proper mental acts. The former has but the latter lacks horizons.

[71] Moreland, J. P. (2002). Naturalism, nominalism, and Husserlian moments. *The Modern Schoolman*, *79*(January/March), 199–216. For a brief summary of Husserl's view of categorial intuition, see Hopp, *Perception and knowledge*, pp. 68–73.

[72] Husserl, E. (1970). *Logical investigations* (2nd ed., J. N. Findlay, Trans.). London: Routledge and Kegan Paul, (hereafter, *LI*); *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy. First book: General introduction to a pure phenomenology*, translated by F. Kersten (The Hague: Nijhoff, 1982) (hereafter, *Ideas I*).

[73] *LI* Vol. I, 332, 337, 340; *LI* Vol. II, 441, 446–447, 453; *Ideas I*, 7–8. Cf. *LI* Vol. II, 773–815.

[74] *LI* Vol. I, 340, 357, 361, 379; *Ideas I*, xxi.

(iii) Compared to ordinary perception of a particular, eidetic intuition is a new mode of apprehension. While founded on ordinary perception, eidetic intuition has its own unique act character, and as it is directed to the moment itself (or, perhaps. the concrete particular), though in an essentially different manner than in ordinary perception, the universal is placed before the subject. The shift of focus and act character from viewing a moment as a particular to intuitively perceiving the universal "in" the moment is like a Gestalt switch.[75]

(iv) The (visual) universal is perceived with the eyes while looking at the moment itself (or, perhaps, the concrete particular) in the proper manner, that is, in such a way that a founded act of eidetic intuition occurs.[76]

(v) The universal is not only in the moment, it is also perceived in the moment and perceived to be in the moment.[77]

Self-Knowledge: For Husserl, the referent of "I" in premise (3) is the pure ego (a.k.a. Transcendent Ego), the irrefutable, secure starting point for First Philosophy.[78] The pure ego is the subject of consciousness and the enduring continuant through a series of mental acts and consciousness-of states. It stands "beyond" the world as that to whom the world is given and is the subjective pole of the subject-object duality. Husserl holds that while the pure ego cannot be sensed (pace Hume), it can be apperceived, actively and directly attended to. As we noted above, in direct *de re* awareness (and reference) that count as knowledge, the intentional object must "show up" or be present. When that intentional object is the self that employs the first-person indexical to self-refer, we have a case of direct *de se* knowledge by acquaintance with the self and direct self-reference. On the basis of such intuitive apprehension (and factors to be addressed shortly), Husserl claims that the pure ego could survive the annihilation of the natural world, including the body. Thus, he would affirm (3).

In sum, an act of *a priori*, modal knowledge will involve the appropriate fulfillment structures and full, direct fulfillment in which the relevant entities constituting modal facts and the truth value of propositions about those modal facts are directly apprehended by way of eidetic or categorial intuition. Fulfillment structures are diachronically unified in that *the same self* who begins with empty intensions, directs the series of subsequent states toward fulfillment, and stands at the end of the series with knowledge due to the thing itself being present.

In the chart in Figure 8.2, let IO stand for any intentional object and let immediate awareness include nonconceptual perception, eidetic and categorial intuition, and the self/soul.

As just noted (see Figure 8.2), an act of *a* priori modal knowledge entails that certain modal entities are involved. But what, exactly, are these modal entities? To answer this, we need to look at that aspect of Husserl's ontology relevant to our question.

Husserl's ontology combines an Aristotelian constituent ontology according to which genuine universals—properties and relations (propositions are structural mental

---

[75] *LI* Vol. I, 337, 340, 357, 389, 390–391.

[76] *LI* Vol. I, 337, 357, 383.

[77] *LI* Vol. I, 337, 339–340, 357, 377, 379; *LI* Vol. II, p. 482; *Ideas I*, 8, 25.

[78] Husserl's views on the "ego(s)" are quite complicated and changed over time. For an exposition of his views, along with reference to key primary sources in Husserl's writings, see Matheson Russell's Husserl: *A guide for the perplexed*, pp. 144–61; and Victor Velarde-Mayol, pp. 70–73.

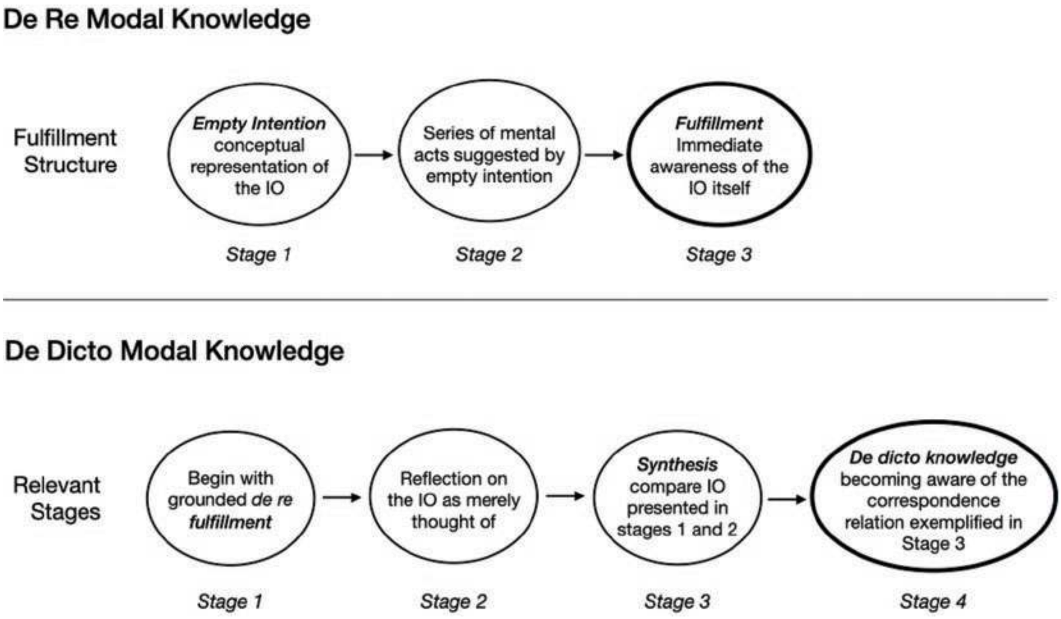

De Re Modal Knowledge
Fulfillment Structure
Empty Intention conceptual representation of the IO
Series of mental acts suggested by empty intention
Fulfillment Immediate awareness of the IO itself
Stage 1
Stage 2
Stage 3
De Dicto Modal Knowledge
Relevant Stages
Begin with grounded de re fulfillment
Reflection on the IO as merely thought of
Synthesis compare IO presented in stages 1 and 2
De dicto knowledge becoming aware of the correspondence relation exemplified in Stage 3
Stage 1
Stage 2
Stage 3
Stage 4

FIGURE 8.2 Husserl's Epistemology of Defeasible Modal Knowledge.

properties)—are immanent within and constitute the species/nature of their metaphysically complex instances (or concrete particulars) with a rich mereological ontology.[79] This ontology provides a crucial backdrop for his modal epistemology, and accordingly, we will describe the former only as it is relevant to the latter.

As noted, Husserl was a realist regarding universals. Besides monadic properties, he held to the existence of what we would call external and internal relations, which we defined in Chapter 3 as follows.

> *Internal Relation*: R is an internal relation between *a* and *b iff* (i) facts about R are grounded in facts about the natures of *a* and *b*, and (ii) necessarily, if R fails to obtain, *a* and *b* are altered.

If the R of *a*R*b* is internal to *a* (or both *a* and *b*), then for all *x*, if *x* does not stand in R to *b*, then $x \neq a$. An internal relation is grounded in the nature/essence of the relevant relatum or relata. Conversely,

> *External Relation*: R is an external relation between *a* and *b iff* (i) facts about R are not grounded in facts about the natures of *a* and *b*; and (ii) possibly R fails to hold between *a* and *b*, while at least *a* or *b* remain unaltered.

If the relation, R, of *a*R*b* is *external to a* (or *a* and *b*) then *a* can exist and retain identity if it fails to stand in R to *b*. Spatial relations, for example, are external.

---

[79] See Moreland, J. P. (2013). Exemplification and constituent realism: A clarification and modest defense. *Axiomathes*, *23*, 247–259.

Recall from (3) that for Husserl, a part is any distinguishable particular "in" a whole.

> *Separable Part*: *x* is a separable part of some whole W *iff* (i) *p* is a particular, (ii) *x* is a part of W, and (iii) it is possible for *x* to exist without being a part of W.
>
> *Inseparable Part*: *x* is an inseparable part of some whole W *iff* (i) *x* is a particular, (ii) *x* is a part of W, and (iii) it is not possible for *x* to exist without being a part of W.

A separable part is independent in the sense that it can exist if the rest of the whole of which it is a part is annihilated. Such parts are accidental to other such parts and their wholes in that they require an external relation to bond together into their wholes (e.g., mereological aggregates of separable parts). Such bonding does not follow from the parts' natures.

Inseparable parts (moment) are dependent entities. They are intrinsically partial in this sense: they require supplementation by something else to exist. That supplemental entity is called the moment's foundation. Barry Smith and Kevin Mulligan offer this sufficient condition for some entity α to be founded on another one β:

> If there is a law of essence that an α as such cannot exist except in a more comprehensive unity which connects it with a β, then we say that an α requires foundation by a β.[80]

Foundation can be one-sided (e.g., an instance of redness cannot exist without foundation in an instance of extension but not conversely) or mutual (an instance of size cannot exist without an instance of shape and conversely). For moments, no external relation is needed for their unification. Rather, their essences require them to be (modally or mutually) unified by way of internal relations into larger wholes.[81]

It is the relations among the essences of various parts that determine the kind of unity various actual/possible wholes do or can have. Husserl's analysis of these wholes, along with the constitutive nature of the essences providing laws of combination among the instances of those essences, is the basis of modal relations.

Husserl distinguishes compatible and incompatible objects. Two objects are compatible if they can be joined together into a whole. And they are incompatible if they cannot be so joined. Compatibility/incompatibility is determined by the immanent essence of the relevant objects, that is, in virtue of *what kind* of object they are, and not by their individuality *per se*. Also relevant to compatibility/incompatibility is the form (essence, type of combinatorial structure) of the whole. If two objects can combine into the unity of a kind of whole, they can do so in part because of the form of unity of the whole. Thus, two objects can be compatible under one form of unity but not another. For example, no physical object can be red and green all over at the same time but can be red and green when these moments are adjacent.

It should be clear that modal concepts are peppered throughout this account. If a set of parts are compatible, relative to a kind of whole, then that whole is metaphysically possible. If the set is incompatible, the whole is impossible. Thus, possibility and impossibility follow from the compatibility/incompatibility of the (usually complex) universals (essences/natures) immanent within the parts of and relative to the form of a kind of whole. *De dicto* modal judgments about possibility/impossibility will be epistemically grounded in eidetic/categorial intuition of the universals and the relevant relations constituting the natures of the parts and the form of the possible whole.

---

[80] Smith, B., & Mulligan, K. (1982). Pieces of a theory. In B. Smith (Ed.), *Parts and moments: Studies in logic and formal ontology* (pp. 15–109). München: Philosophia Verlag; the quote is on 43.

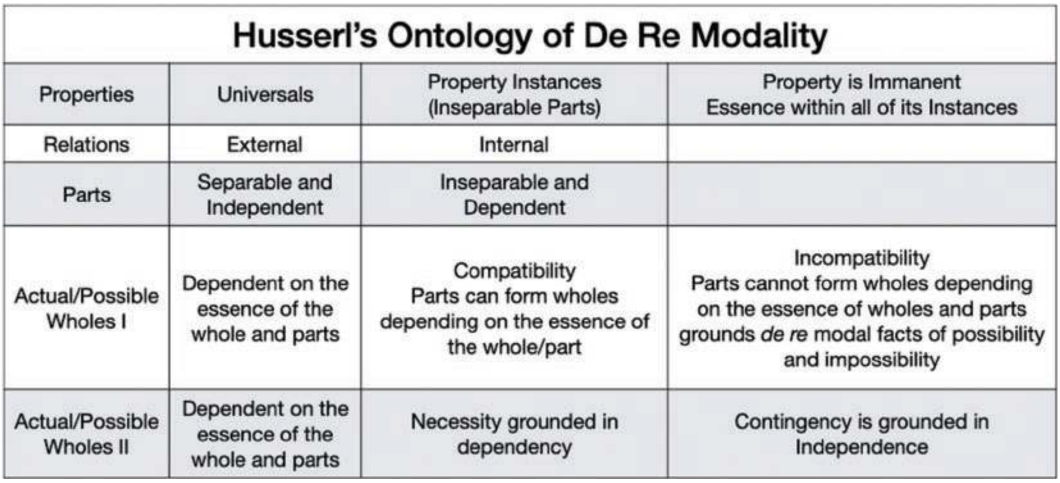
[81] See Moreland, J. P. (1989). Was Husserl a nominalist? *Philosophy and Phenomenological Research*, *49*(June), 661–674.

Further, modal knowledge involves the ontology of dependence and necessity. The universal essences and internal relations prescribe to their actual/possible instances their relative dependencies that serve as the ultimate ground of modality. For example, necessity and contingency are grounded in relations of dependence (one-way or mutual) and independence, respectively. This provides an analysis of the entities that constitute modal facts and for which eidetic intuition grounds modal knowledge. Modal intuitions begin with eidetic intuitions of universals, and they ground the further discovery of the dependency and compatibility relations that ground various modalities and are the objects of modal knowledge.

Figure 8.3 summarizes the main items in Husserl's ontology of *de re* modal facts.

Since Husserl's views are not widely known, and given space considerations, we cannot provide an account of sources of modal error (it will involve problems with intuitive attentiveness, fulfillment structures that go off course, etc.) and answers to objections raised against his views.[82] However, Figure 8.3 provides a sufficient summary of Husserl's views for our purposes.

In this section, while Husserl's account of modal knowledge is close to our own, our main purpose has been to provide a range of options regarding modal knowledge and error as part of the defense of (3). We now turn to a brief consideration of several objections against modal knowledge, premise (3), and/or the modal argument in general.



Husserl's Ontology of De Re Modality

Properties	Universals	Property Instances (Inseparable Parts)	Property is Immanent Essence within all of its Instances
Relations	External	Internal	
Parts	Separable and Independent	Inseparable and Dependent	
Actual/Possible Wholes I	Dependent on the essence of the whole and parts	Compatibility Parts can form wholes depending on the essence of the whole/part	Incompatibility Parts cannot form wholes depending on the essence of wholes and parts grounds *de re* modal facts of possibility and impossibility
Actual/Possible Wholes II	Dependent on the essence of the whole and parts	Necessity grounded in dependency	Contingency is grounded in Independence

Figure 8.3 Husserl's Ontology of *de re* Modality.

---

[82] See Kasmier, chapters 7 and 8.

8.3 FIVE OBJECTIONS AGAINST MODAL ARGUMENTS FOR SD[83]

8.3.1 The Modal Argument Is Circular

Stewart Goetz has argued that the modal argument is viciously circular in this way: Knowledge of a premise like C_5 (I have the essential property of being such that I possibly survive disembodied (without my or any other physical body)) presupposes knowledge of C_4 (I am not identical with my (or any) physical body[84]). Knowledge of C_5 requires I already know C_4 which, in turn, is likely due to an awareness of a real distinction between me and my body. In turn, this is due to a further awareness of a difference in their essential properties (e.g., primarily, being simple vs. being complex, but also, being unextended vs. being extended, being constituted by psychological properties vs. being constituted by physical properties). If we reject this claim, then some sort of conceivability will be the only thing available to the modal dualist for C_5, and it will most likely be weak conceivability (failing to be aware of the necessary connection between me and my body), which is inadequate to support C_5 since an absence of awareness is not awareness of absence.

Reply1. We agree with Goetz that we have direct awareness of our own I and body (or other physical objects), and on that basis, we are aware of a genuine metaphysical distinction between them. But Goetz's argument assumes direct awareness of the self (and its simplicity and unextendedness), and many philosophers reject such a thing. So, even if Goetz is correct, we see plenty of room for a modal argument for those who reject Goetz's approach.

Reply 2. Charles Taliaferro responds by claiming that we have strong conceivability plus the absence of epistemic negligence (roughly, inadvertently careless awareness or reflection) with respect to the real distinction expressed in C_5) and this is adequate grounds for $C_{5.}$[85] Goetz responds that at best, Taliaferro's rejoinder supports property dualism because his strong conceivability targets the properties involved in our contingency. Moreover, Taliaferro never specifies what properties these are, and if they are psychological ones, property dualism is all Taliaferro can justify. The I and contingency are left out.

It may be that Taliaferro does not adequately specify the relevant properties. We don't have an opinion on that. But we think Husserl did. For Husserl, we have intuitive awareness of the I and an eidetic intuition of its essence (e.g., being simple with respect to separable parts; having mental properties or the capacities for them), we have categorial intuitive awareness of the subcategory "Material Object," and on that basis, we become aware of the independence and separability of the two. Husserl may be wrong, but this must be shown if his way out is to be rejected.

---

[83] We are omitting a debate between William Hasker and Richard Swinburne about whether or not Swinburne's version of the modal argument is circular (Jaegwon Kim actually takes it to be an argument about persistence conditions). See Hasker, W. (1998). Swinburne's modal argument for dualism: Epistemically circular. *Faith and Philosophy, 15*, 366–370; Swinburne, R. (1998). The modal argument is not circular. *Faith and Philosophy, 15*, 371–372. We are also setting aside the view that one can accept all the premises of a modal argument and deny the conclusion advanced by advocates of material constitution physicalism. See Taliaferro, C. (2009). *Philosophy of religion* (pp. 126–127). Oxford: Oneworld Publications.

[84] Goetz, S. (2001). Modal dualism: A critique. In K. Corcoran (Ed.), *Soul, body and survival* (pp. 89–104). Ithaca, NY: Cornell University Press.

[85] Taliaferro, C. (1986). A modal argument for substance dualism. *The Southern Journal of Philosophy, XXIV*, 95–108.

8.3.2 The Modal Argument Establishes Merely a Duality of Concepts

Jaegwon Kim is among a number of philosophers who have raised this objection.[86] The basic idea is that in certain intensional contexts, especially those involving *de dicto* beliefs, equals can be substituted for equals, but the truth value is not preserved. Thus, the man who sped away is believed by the police to be a hit-and-run driver, the man who sped away is, in fact, my neighbor, but the police do not believe my neighbor is the alleged driver. Here we have two concepts that refer to the same object, not two different objects, even though it surely seems possible that there are truths about my neighbor that are not true of the man who sped away. And the same problem is exhibited by the modal argument.

Reply 1. Our modal argument employs rigid designators in (3), (4), (5); they involve *de re* beliefs and are epistemically grounded in *de re* knowledge by acquaintance. Kim-style counterarguments exhibit *de dicto* beliefs (and, often, nonrigid designators). So the counterargument is importantly dis-analogous to the relevant premises in our modal argument. To see this, substitute *de re* beliefs for the *de dicto* beliefs in his counterargument. We now have that the police believe of that man there who sped away (namely, my neighbor) that he is the hit-and-run driver, but the police do not believe of the man who is, in fact, my neighbor that he is the hit-and-run driver. Here, the second belief is false because the police do, in fact, believe *of* my neighbor that he is the driver. They just don't believe *that* he is my neighbor.

Reply 2. It is clear that the problem surfaced in Kim-style arguments is based on a lack of knowledge of something relevant. And it's easy to see how one could have that lack of knowledge, e.g., that the police would have no idea that the speeding driver is my neighbor. But it isn't easy to see where the source of error is in our modal argument. If we do, in fact, have direct awareness of our I and find it to be essentially simple, non-extended, and constituted by mental properties, and if we have a pretty good third-person set of descriptions of physical objects, including one's body, where does the error come from? Husserl's account gives a detailed analysis of the sort of entities known that are relevant to the modal argument, and we see no non-question-begging identification of modal error involved in the argument.

8.3.3 Physicalist and Dualism Intuitions Cancel Out Each Other

Peter van Inwagen has argued that thought experiments with reverse modal or physicalist intuitions are easily constructed and cancel out dualist intuitions.[87] For example, van Inwagen claims that "... when I enter most deeply into that which I call *myself*, I *seem* to discover that I am a living animal. And, therefore, dualism seems to me to be an unnecessarily complicated theory about my nature."[88]

---

[86] Kim, J. (2006). *Philosophy of mind* (2nd ed., pp. 35–36). Boulder, CO: Westview Press.

[87] van Inwagen, P. (1998). Modal epistemology. *Philosophical studies*, *92*, 67–84. For a response, see Taliaferro, C. (2001). *Sensibility and Possibilia*: A defense of thought experiments. *Philosophia Christi, 3*(2), 403–420.

[88] van Inwagen, P. (1995). Dualism and materialism: Athens and Jerusalem? *Faith and Philosophy*, *12*(4), 475–488. The quote comes from p. 476.

Reply 1. We doubt that such intuitions are actually possible, and if people have them, we would agree with Husserl that an inadequate concept of I and of a material body prevent fulfillment when the things themselves are given directly to intuitive awareness (which can occur with actual and merely possible entities). In our view, physicalists may have physicalist beliefs, but we suspect they retain dualist intuitions. But setting that aside, even if such intuitions are possible, they are overwhelmed by dualist intuitions. Here's why.

As we argued in Chapter 5, almost everyone acknowledges that we all have, or at least start with, dualist intuitions. But this is not a throw-away line. It expresses overriding epistemic justification. So it is false that physicalist intuitions are sufficiently strong to cancel out dualist ones.

Reply 2. Even if we grant that physicalist/dualist intuitions cancel out, it is not clear what is supposed to follow. The growing literature in the epistemology of peer disagreement exemplifies disagreement about what the demands of rationality require when such disagreement occurs. Moreover, for those who adopt some version of the Steadfast View, one may be justified or warranted in continuing to hold to one's view even when peer disagreement seems irresolvable.

8.3.4 Problems with Dualist Thought Experiments

According to van Inwagen, while we have modal knowledge in clear, everyday cases, we have no clear conception of how a thought experiment might work to provide credible views of what is possible, especially regarding more sophisticated philosophical claims. Moreover, presenting stories or detailed scenarios doesn't tell us how we know the state of affairs is possible. And such scenarios would have to be presented at such a high level of detail/comprehensiveness that they would surpass our capacities.[89]

Reply 1. Though defeasible, thought experiments are highly successful throughout philosophy for focusing our intuitive attention of specific items. We are not claiming merely that because this is all we have, we just go with it. Rather, we are claiming that their use in focusing/clarifying modal seemings is demonstrably *successful* and, regarding the modal argument, since dualist intuitions conform to those of folk ontology, *sans* a question-begging commitment to physicalism, we see no reason to deny their use in the modal argument.

Reply 2. Van Inwagen accepts modal intuitions in clear, commonsense cases, but rejects them in more sophisticated philosophical cases. But dualist intuitions have been pervasive throughout the world for millennia, and very young children pervasively have them. All the modal argument does is to capture and focus these ubiquitous, commonsense cases because a lot of philosophers don't like where they lead. So the intuitions captured in the modal argument are clear and common sensical.

Reply 3. You can know something, especially something basic and fundamental like dualist intuitions, without knowing how you know them. Moreover, the purpose of a thought experiment expressed as a scenario is to add a deeper focused awareness to surface intuitions that we know such and such; it is not to show how we have such knowledge. That is the task of developing theories of modal knowledge.

---

[89] See van Inwagen, P. (1998). Modal epistemology. *Philosophical Studies*, *92*, 67–84.

Reply 4. It is not true that we don't have an account of how we have the modal knowledge relevant to the modal argument. We have presented three such accounts, and with slight adjustments, O'Connor's and Bealer's views can easily accommodate the intuitions relevant to the modal argument. And Husserl's view is a detailed, rigorous account of modal knowledge and how we have it, including such knowledge in the modal argument.

Reply 5. It is not true that adequate thought experiments require a level of detail that surpasses our cognitive capacities. There is a vast and sophisticated body of literature on near death experiences (NDE). In that literature are literally thousands of highly detailed, comprehensive accounts of disembodied existence, far more so than successful thought experiments in other areas of philosophy. Upon examination, one becomes aware of their adequacy of detail/comprehensiveness, along with the intuitive insight that NDE accounts contain no incompossible states of affairs or other entities. They are metaphysically possible.

8.3.5 Kripke and the Confusion Between Epistemic and Metaphysical Possibility

In most cases, a claim about the link between what seems possible and what is possible is defeasible. Accordingly, a widely employed criticism of the modal argument is that it establishes a mere epistemic, not a metaphysical, possibility in (3). This criticism was strengthened by and currently, usually expressed in association with Kripke's arguments in support of *a posteriori* necessities.

Reply. It would take an entire book to address this issue adequately and, due to space considerations, we can only gesture at two lines of response. First, Richard Swinburne has developed an account of Informative Rigid Designators that, if successful, solves this problem by showing that the designators in the modal argument are informative ones and, thus, our modal knowledge of (3) is *a priori*.[90] David Chalmers (and Frank Jackson) have developed the widely known solution called two-dimensional semantics which, if successful, show the relevant necessities, at least in Kripkean cases, are *a priori*.[91]

Second, Kripke's argument about metaphysical/epistemic confusion rests on the relevant community failing to know something. But the accounts of modal knowledge presented above, especially Husserl's, show that the modal intuitions involved in the modal argument are grounded in positive intuitive awareness of the relevant entities and their metaphysical modality. Thus, Kripke's cases are dis-analogous to the modal argument.

In this chapter, we have formulated and defended a modal argument for SD. We acknowledge that, granting the success of my modal argument, it doesn't follow that SD is true or the best philosophical anthropology to adopt. That judgment would need to be done in light of all of the relevant considerations, and the modal argument is just one of them. But it is an important one, and in any case, we trust this chapter will provide resources for further dialog.[92]

---

[90] Swinburne, R. (2013). *Mind, brain and freewill* (pp. 11–14, 23–24, 27–28, 36–38). Oxford: Oxford University Press.

[91] Chalmers, D. J. Does conceivability entail possibility? In *Conceivability and possibility* (pp. 145–200).

[92] We are grateful to the following people for their helpful feedback on various parts of this chapter: Andrew Bailey, Stewart Goetz, Walter Hopp, Ross Inman, Jonathan Metcalf, Joshua Rasmussen, and Charles Taliaferro.

CHAPTER 9

Staunch Libertarian Agency and the Simple, Enduring Soul

In his excellent defense of agent causation, *Persons and Causes*, Timothy O'Connor lays out his project in this way:

> Philosophers want to do more than paint a commonly held picture of ourselves. We also want to put forth a vision of human beings and their place in the wider scheme of things … There is at present a widespread trend in English-speaking philosophy toward 'deflationary' analyses of most traditional targets of philosophical inquiry … The general, if vague, impetus is to analyze philosophical notions in a way that makes them hospitable to a 'naturalistic' view of human beings that has apparently been handed down to us by 'Science'.[1]

In context, O'Connor's remarks are applied to questions about human action. His approach is to counter the "fashion for apologies on behalf of Naturalism," by, among other things, giving pride of place to pre-philosophical intuitions and philosophical arguments about human agency. Our project in this chapter includes O'Connor's, even though we depart from his conclusions at certain points. In doing so, we advance an argument for substance dualism from certain features of free will.

---

[1] O'Connor, T. (2000) *Persons and causes* (pp. xi–xii). NY: Oxford University Press.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.

© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

9.1 CONTEMPORARY RECOGNITION OF THE PROBLEM SLA RAISES FOR NATURALISM AND PHYSICALISM

It is widely acknowledged that worldwide, commonsense, spontaneously formed understanding of human free will is what philosophers call libertarian freedom[2]: One acts freely only if one's action was not determined—directly or indirectly—by forces outside one's control, and one must be free to act or refrain from acting; one's choice is "spontaneous," it originates as a fundamental exercise of active power with the agent, and it is often done for reasons construed as teleological ends.

We shall call this view *Staunch Libertarian Agency* (SLA) and unpack it below. Since we will be guided by the ontological commitments of Folk Ontology,[3] in Chisholm's sense, we take ourselves to have a right to be sure about SLA and believe a considerable burden of proof exists for those who water it down or outright reject it.[4]

We believe a phenomenologically and dialectically adequate account of free will must do the following:

1. Provide a *depiction of the agent* that adequately grounds the agent's ability in non-basic acts to sustain, guide, and adjust steps in an action plan toward an intended end so as to stand at that end as a responsible agent.
2. Provide an *account of control* that grounds genuine responsibility for acting and provides a non-arbitrary, regress-stopping primitive.
3. Provide an account of free choice that *makes clear the reality and nature of alternative possibilities* at the time of choice while avoiding the problem(s) of luck.
4. Provide a *non-deterministic account of reasons* that explains their role in guiding and, if needed, redirecting a series of choices in an action plan done on account of that reason.

What is also widely acknowledged is that staunch naturalism and faint-hearted naturalism, strong and weak physicalism are inconsistent with SLA. Thus, it is evident to most analytic philosophers who are philosophical naturalists that their ontological framework rules out, or at least, makes highly implausible, the reality of libertarian free will and agency. As Kevin Timpe and Jonathan D. Jacobs correctly observe: "There is ... a near consensus on this: naturalism is certainly incompatible with *libertarian* free will."[5] Accordingly, Roderick Chisholm claimed that "... in one very strict sense of the terms, there can be no science of man."[6] Along similar lines, John Searle says that "our conception of physical reality simply does not allow for radical [libertarian] freedom."[7] And though a bit dated, in what still may be the best naturalist attempt to reconcile naturalism and the sort of freedom (he opts for compatibilism) necessary for moral responsibility, John Bishop frankly admits that

---

[2] For an insightful, brief description of the way that naturalism brings about the "disappearance of the agent," see Lowe, E. J. (2008). *Personal agency* (pp. 159–62). Oxford: Oxford University Press.

[3] Physicalists, such as Paul Churchland, commonly reject folk ontology. For an excellent, yet overlooked, treatment of Paul Churchland's rejection of folk psychology in relationship to folk ontology, see Haldane, J. (1988). Understanding folk. *Proceedings of the Aristotelian Society*, Supplementary, 62, 223–254.

[4] According to Chisholm, if we have the (epistemic) right to be sure of some proposition P, then this right does not entail termination of inquiry but, rather, is the claim that we are entitled to take ourselves to (defeasibly) know P and to rely on P to do work and judge other things in philosophy. See his *Theory of knowledge* (2nd ed., pp. 116–118). (1966). Englewood Cliffs, NJ: Prentice-Hall.

[5] Timpe, K., & Jacobs, J. D. (2016). Free will and naturalism: How to be a libertarian, and a naturalist too. In K. J. Clark (Ed.). *The Blackwell companion to naturalism* (p. 319). Malden, MA: John Wiley and Sons, Inc.

> the idea of a responsible agent, with the 'originative' ability to initiate events in the natural world, does not sit easily with the idea of [an agent as] a natural organism … .Our scientific understanding of human behavior seems to be in tension with a presupposition of the ethical stance we adopt toward it.[8]

Nevertheless, some demur and believe there are versions of naturalism and libertarianism that are consistent. To cite one example, Timpe and Jacobs state: "We aim to show in this paper that libertarian free will is not incompatible with naturalism."[9] And those who agree with the inconsistency of the two often fail to provide a general ontological framework for why they think this. As an aside and for what it's worth, we think that too often this failure results from either a rejection of first philosophy or a limitation of it to mere conceptual analysis.

These ubiquitous claims cited above have led many naturalists to agree with Leda Cosmides and John Tooby, who recognize that epiphenomenalism follows from faint-hearted naturalism:

> The brain is a physical system whose operation is governed solely by the laws of chemistry and physics. What does this mean? It means that all of your thoughts and hopes and dreams and feelings are produced by chemical reactions going on in your head.[10]

Finally, James Rachels reminds us that naturalistic evolutionary theory alone undermines the reality of free will:

> aspects of our behavior which we previously thought were matters of free choice are really the products of deep, genetically controlled forces … Humans may fabricate all sorts of other reasons for what they do, but these are mere rationalizations. In reality their conduct is fixed by the territorial imperative built into their genes.[11]

Accordingly, our purpose in this chapter is to show that (i) staunch naturalism and strong physicalism render SLA virtually impossible; (ii) faint-hearted naturalism/weak physicalism do not provide plausible grounds for embracing SLA that render it consistent with these ontological commitments, and they are inordinately guided by a naturalistic worldview in contradistinction to the approach proffered by O'Connor and adopted by us; (iii) representative alternatives to SLA are inadequate for numerous reasons. Granting that we have SLA and that a range of ontological commitments constitute it, it follows that we are simple enduring souls; (iv) as the quote from Cosmides and Tooby illustrates, setting aside libertarian freedom, there are a number of difficulties with mental causation in general (e.g., the causal closure of the physical, difficulties with causal interaction). While we will address some of these difficulties in this chapter and in Chapter 10, our primary focus here is on SLA.

---

[6] Chisholm, R. M. (1989). Human freedom and the self. reprinted in *On Metaphysics* (p. 14). Minneapolis, MN: University of Minnesota Press.

[7] Searle, J. (1984). *Minds, brains, and science* (p. 98). Cambridge, MA: Harvard University Press.

[8] Bishop, J. (1989). *Natural agency* (p. 1). Cambridge: Cambridge University Press. Bishop's own solution eschews libertarian agency in favor of a version of compatibilism.

[9] Timpe and Jacobs, p. 319.

[10] Cosmides, L., & Tooby, J. (1997). *Evolutionary psychology: A primer* (p. 4). University of California at Santa Barbara: Center for Evolutionary Psychology. https://scholar.google.com/scholar?q=Leda+Cosmides+and+John+Tooby,+Evolutionary+Psychology:++A+Primer&hl=en&as_sdt=0&as_vis=1&oi=scholart.

[11] Rachels, J. (1990). *Created from animals* (pp. 73–74). Oxford: Oxford University Press.

We acknowledge that this focus limits the dialectical success of our argument to those who embrace libertarian freedom. However, since more general difficulties with mental causation have been debated over and over again, and since, relatively speaking, less has been said regarding naturalism and SLA, we wish to fill this lacuna and hope the chapter will be of use to those who reject any version of libertarian free will.

9.2 THE ARGUMENT AND ITS BASIC EPISTEMIC JUSTIFICATION

As previously noted, in laying out our case, we will be guided by this advice from Joshua Hoffman and Gary S. Rosenkrantz:

> If entities of a certain kind belong to folk ontology [the ontological presuppositions of our common sense conceptual scheme], then there is a *prima facie* presumption in favor of their reality. ... [T]hose who deny their existence assume the burden of proof.[12]

9.2.1 The SLA Argument

Consider the following, what we will call the SLA Argument:

1. If human persons are wholly physical objects (e.g., a brain or a body), then they do not have SLA.
2. But human persons do have SLA.
3. Therefore, human persons are not wholly physical objects.
4. Human persons either are wholly physical objects, or they are (or have) simple enduring souls.
5. Therefore, human persons are (or have) simple enduring souls.

9.2.2 Our Fundamental Justification for Libertarian Freedom

Most philosophers working in the subdiscipline regarding action theory and agency approach questions of free will by focusing on the sort of control needed to underlie moral responsibility. We agree that this is important (though we would add rational control to this approach), but we do not think that this is the conceptual or epistemic starting point for developing a view of agency, especially a libertarian one. The control approach leaves open the question as to why people have a concept of moral responsibility and a belief that they are morally responsible agents in the first place. Further, the control approach can, and sometimes does, limit (libertarian) free action to moral acts, especially momentous character-forming ones. But this is a mistake. The primary places we exercise free agency are in directing our attentive noticing—our intentionality—toward various intentional objects and in moving our bodies. As one goes through one's day, one freely chooses to think, perceive, and notice various things within or around him/her.

---

[12] Hoffman, J., & Rosenkrantz, G. S. (1997). *Substance: Its nature and existence*. (p. 7). London: Routledge. Cf. 77–79.

How does one know that he/she has free agency? In our view, it is fundamentally through knowledge by acquaintance with one's own states of consciousness (and with their being one's own)—especially volitional acts—from within the first-person point of view.[13] It is through such direct introspective seeming that people become aware of and accept (or should accept unless they have an ideology that disallows) their possession and exercise of active power. A basic exercise of active power has what Carl Ginet calls an "actish phenomenal quality," a "seeming directly to produce" that is distinct from the experience of a passive liability being triggered, and both are available to direct awareness.[14] Like the dualist seemings analyzed in Chapter 5, libertarian free will seems obvious to almost all human persons. Moreover, there is good evidence that this free will intuition is grounded in one's direct self-awareness of self, one's thoughts and actions.[15] Research also shows a correlation between belief in free will and belief in generic substance dualism (SD).[16]

We also seem to be aware directly of the phenomenological difference between deciding and acting or, as some put it, between the final-stage summing-up judgment (e.g., "Given all the relevant factors, doing A is preferable to doing B") and acting or the final-stage forming of a pure intention to act (e.g., "let A be the case over B") and acting itself.[17] This awareness is one of the reasons for justifying final stage *akrasia*, a "weakness of will" that involves acting contrary to one's overall summing-up judgment or one's final-stage formed pure intention that occurs between these states and the act itself.

---

[13] Though his topic is sensory perception, Walter Hopp argues persuasively that perceptual acts are constituted by two kinds of non-conceptual content: horizontal and intuitive, and that these direct a perceptual act upon its intentional "object." With modifications for direct awareness of one's own mental states, we are extending Hopp's account with which we agree. See Hopp, W. (2011). *Perception and knowledge* (pp. 54–73, 149–189). Cambridge: Cambridge University Press.

[14] Ginet, C. (1990). *On action* (p. 13). Cambridge: Cambridge University Press.

[15] See, e.g., Leotti, L. A., et al. (2010). Born to choose: The origins and value of the need for control. *Trends in Cognitive Sciences*, *14*, 457–463; and Sarkissian, H., et al. (2010). Is belief in free will a cultural universal? *Mind and Language, 25*, 346–358. Cf., Forstmanna, M., & Burgmerb, P. (2018). A free will needs a free mind: Belief in substance dualism and reductive physicalism differentially predict belief in free will and determinism. *Consciousness and Cognition*, *63*, 280–293.

[16] See, e.g., Fernandez-Duque, & Schwartz, D. (2016). Common sense beliefs about the central self, moral character, and the brain. *Frontiers in Psychology*, *6*(article 2007), 1–12; Nadelhoffer, T., et al. (2014). The free will inventory: Measuring beliefs about agency and responsibility. *Consciousness and Cognition*, *25*, 27–41; and Nadelhoffer, T., & Tocchetto, D. G. (2013). The potential dark side of believing in free will (and related concepts). In Caruso, G. (Ed.), Exploring the illusion of free will and moral responsibility (pp. 121–141). Lanham, MD: Lexington Books; Nadelhoffer, T., Murray, S., & Murry, E. (forthcoming). Intuitions about free will and the failure to comprehend determinism. *Erkenntnis*. doi: 10.1007/s10670-021–00465-y; Vargas, M. (2014). Reconsidering scientific threats to free will. In W. Sinnott-Armstrong (Ed.), *Moral psychology*, Vol. 4: *Free will and moral responsibility* (pp. 417–425). MIT Press; and Wisniewski, D., Deutschländer, R., & Haynes, J.-D. (2019). Free will beliefs are better predicted by dualism than determinism beliefs across different cultures. *PLoS ONE*, *14*(9), 1–19.

Consider this thought experiment. Most of us are quite aware of the different what-it-is-like, the different phenomenological texture in having a passive thought (e.g., when someone is talking to me) and an active thought (one in which one exercises one's active power and chooses to have).[18] This difference is self-evident to introspective awareness. And such awarenesses provide non-doxastic, internalist grounds for the proper basicality of one's belief that one has and can exercise active power. Thus, one's relevant introspective awareness of active power is epistemically and conceptually foundational and essential for there to be such things as a concept and knowledge of what it is to exercise active power, along with knowledge that one actually exercises such power. And it is on the basis of this account and knowledge of relevant moral features that we have the concept of moral responsibility.

Stephen Mumford has developed a metaphysical and epistemic account of our direct awareness of our own powers (which he uses interchangeably with "dispositions"). Mumford argues that causation construed as an actualization or exercise of powers can be directly "perceived" especially when we "consider human agency. We are after all causal agents (and patients) ourselves, and so we have some first-hand experience of powers at work, namely our own."[19] Through a certain aspect of proprioception, says Mumford, we have direct awareness or immediate knowledge of our own causal powers actively exercised or passively triggered.

We think that these observations explain why a libertarian concept of free will and of the exercise of active power entailed by SLA are the parameters within which philosophers develop accounts—including compatibilist ones—of free will. To cite a clear case of this, John Bishop develops a rigorous, detailed account of compatibilist free will that relies on an event-causal theory of action constituted by appropriate chains of triggered passive liabilities.[20] His methodology for achieving his settled account is controlled from beginning to end by a commonsense libertarian concept of free agency and active power.

Bishop identifies his goal to be the formulation of a compatibilist theory that approximates a libertarian view as closely as possible. Thus, he starts his project by clarifying what he takes to be the commonsense concept of libertarian free will, formulates a conception of compatibilist freedom that is close to that concept, raises counterexamples to the formulation that are derived from the libertarian conception, adjusts his version of compatibilism to solve the counterexamples by using the libertarian concept to judge the adequacy of his adjusted formulation, and repeats this epicycle until he lands on his final position.

---

[17] See Moreland, J. P., & Rae, S. (2000). *Body and soul*: *Human persons and the crisis in* ethics (pp. 138–140). Downers Grove, IL.

[18] While our argument focuses on the phenomenological difference between an active choice and a passive triggering, for an excellent defense of the idea that cognitive acts like thinking events have unique phenomenal qualia, see Pitt, D. (2004). The phenomenology of cognition or *What is it like to think that P*? *Philosophy and Phenomenological Research*, *LXIX.1*(July), 1–36.

[19] Mumford, S. (2013). The power of power. In R. Groff & J. Greco (Eds.), *Powers and capacities in philosophy: The new Aristotelianism* (New York: Routledge, 9–24). The quote is on page 19.

[20] John Bishop, *Natural agency*.

The entire dialectical process is guided by the relevant libertarian concepts. Why is this so often the case? Because, says Bishop, we all start with a libertarian concept of free will. But why is that the case? One could offer an evolutionary just-so story about why the ubiquitous presence of such a concept provided us with enhanced survival value. But such a story explains only the utility of a libertarian concept once it is formed. However, it fails to explain how such a concept arose in the first place. Further, this story is precisely the sort of account O'Connor warned us about employing. In a moment of revealing candor, Thomas Nagel correctly noted:[21]

> My guess is that this cosmic authority problem [roughly, the desire that no God to whom one is accountable actually exists] is not a rare condition and that it is responsible for much of the scientism and reductionism of our time. One of the tendencies it supports is the ludicrous overuse of evolutionary biology to explain everything about life, including everything about the human mind.

We think a much better answer is staring us in the face. The reason the commonsense understanding of free will is libertarian and the reason this understanding is so widely employed in doing work in the literature on free will is that we all have direct awareness of exercising active power and on that basis, come to realize that it entails the presence of alternative possibilities the selection of which is up to the agent.

It can be objected that such introspective awareness of the "actish feel" of an exercise of active power is significantly defeasible and, in fact, a mental event with such a phenomenological texture could actually be passively brought about by a direct stimulation of one's brain. We offer two responses to this objection. First, from the fact that such introspective awareness *might* be mistaken, it does not follow that on a given occasion, we *have good reasons to think we are, in fact, mistaken*. The skeptic must do more than suggest the mere possibility of error with respect to numerous apparently veridical cases where we seem to be directly aware of an exercise of active power and a triggering of passive liability. We do not accept a burden of proof in this matter.

Second, still relevant and affirmatively employed are data from the electrical stimulatory exploration of the brain surface collected by Wilder Penfield during awake surgery for the treatment of epilepsy. Motor cortex stimulation would cause the hand of the patient to move. Patients would respond: "I didn't do that, you did." By contrast, when a movement was brought about by the patient's own exercise of active power, the patient reported having freely willed the movement.[22] So it seems that, in fact, patients are often, perhaps usually, able to make accurate judgments based on introspective awareness.

9.3 SLA AND ITS METAPHYSICAL WORLDVIEW IMPLICATIONS

According to O'Connor, there are three main incompatibilist approaches to free will: the uncaused-volitional view (Stewart Goetz, E. J. Lowe), the non-deterministically caused events view (Alfred Mele, Robert Kane), and the agent/substance-caused events view

---

[21] Nagel, T. (1997). *The last word* (p. 131). New York, NY: Oxford University Press.

[22] See Penfield, W. (1975). *The mystery of the mind: A critical study of consciousness and the human brain*. Princeton, NJ: Princeton University Press.

(Randolph Clarke, Timothy O'Connor.) We think the view most consistent with Folk Ontology is a version of the agent-causal position we call SLA. We will examine Robert Kane's view below as a paradigm case of the non-deterministically caused events view. We will set aside consideration of the uncaused-volitional view, but we should note that certain ontological commitments usually associated with this approach (a substantial, enduring self, irreducible teleology, exercises of active power) present difficulties with staunch and faint-hearted naturalism. We will discuss these commitments in the context of examining SLA. Still, our approach is not meant to imply that we do not think the uncaused view is a viable position.

9.3.1 A Formal Characterization of SLA

Stated formally,

> *Strong Libertarian Agency* (SLA): A person P exercises SLA and freely does some intentional act *e* just in case (i) P is an enduring substance that has the active power to bring about *e*; (ii) P exerted his/her active power as a first, unmoved mover (an "originator") to bring about *e*; (iii) P had the categorical ability to refrain from exerting his/her power to bring about *e*; (iv) P acted for the sake of reasons which serve as the final cause or teleological goal for which P acted.

Taken alone, (i)–(iii) state necessary and sufficient conditions for a pure voluntary act, e.g., freely directing my eyes toward a specific desk upon entering a room. Propositions (i)–(iv) state necessary and sufficient conditions for an intentional act, i.e., a voluntary act done for a reason (e.g., raising my hand to vote).

9.3.2 Six Features That Constitute SLA

There are six features of a free act that make it difficult and, indeed, virtually impossible to reconcile with the naturalist standpoint.

9.3.2.1 Free Agents Are Substances

First, the free agent is a substance and not an event, a bundle of events, a mereological aggregate, or dynamic system. But for two reasons, ordinary objects are not substances in a naturalist ontology. For one thing, according to naturalism, all causes and effects are events (or bundles of events). The laws of nature govern causal processes in which a temporal state (an event) of an object (an electron, water molecule, storm cloud) brings about a different temporal state (a subsequent event) according to a natural law. This is called "event–event causation," it is constituted by and only by triggering of passive liabilities formed into chains in keeping with natural laws, and it is the only sort of efficient cause recognized in physics, chemistry, geology, biology, neuroscience, or other naturalistically certified hard sciences. Further, such causation is the paradigm case of transuent causation. Event–event causation governs changes of state in or among objects—nothing more, nothing less.

Substances as substances—essentially characterized particulars, substantial things—do not cause things or perform actions. Strictly speaking, it is not the first billiard ball that moves the second. It is the moving-of-the-first-ball (the causal event) that causes the moving-of-the-second-ball (the effect event). Thus, whatever there is to an “object” that is more than its event status is otiose as far as causality is concerned, and Ockham’s razor and the causal criterion of being (something exists only if it can serve as a cause) have led many naturalists to shave off that excess baggage and reduce macro-objects to bundles of events.

For another thing, the Grand Story tells a tale about how macro-objects come-to-be, and that combinatorial tale implies that macro-objects are mere aggregates or systems, relational collections of parts standing in external relations to each other, and not genuine substances (see Chapters 2 and 7). As such, the relevant physical macro-objects are not enduring continuants.

This second point provides a defeater to the following objection: A small but growing number of philosophers are advancing the view that substance-causation is fundamental, and event-causation is derivative and reducible to substance-causation, so the latter is not as odd as it used to be thought.[23] Unfortunately, most of the entities employed in advocating substance-causation are not substances at all. They are mereological systems, aggregates or masses of stuff bereft of teleology and subject to mereologically essentialist constraints.

By contrast, it is the agent as a substantial simple self or I, not some state in the agent, that brings about a free act. As defined and used in previous chapters, by “substance” we mean a member of a natural kind, an essentially characterized particular that sustains absolute sameness through (accidental) change and that possesses a primitive unity of inseparable parts, properties, and capacities/powers at a time. This strong view of substance is required for SLA for at least three reasons:

1. Libertarian agency is possible only if there is a distinction between the capacity to act or refrain from acting and the agent that possesses those capacities.
2. The type of unity present among the various capacities possessed by an agent is the type of unity (i.e., a diversity of capacities internally related to and within an ontologically prior whole) that is entailed by the classic Aristotelian notion of substance.
3. Non-basic free acts take time and include sub-acts as parts, and an enduring agent is what gives unity to such acts by being the same self who is present at the beginning of the action as intentional agent, during the act as teleological guider of means to ends, and at the end as responsible actor. Enduring substantial agents “originate” and sustain free actions that constitute action plans in light of their own ends.

But this view of human persons is not countenanced by staunch naturalism nor is it plausible, given faint-hearted naturalism. Thus, staunch naturalist John Bishop frankly admits,

> the problem of natural agency is an ontological problem — a problem about whether the existence of actions can be admitted within a natural scientific perspective … [A]gent causal-relations do not belong to the ontology of the natural perspective. Naturalism does not essentially employ the concept of a causal relation whose first member is in the category of person or agent (or even, for that

---

[23] See Lowe, J. (2008). *Personal agency* (pp. 141–149, 161–165). Oxford: Oxford University Press.

matter, in the broader category of continuant or 'substance'). All natural causal relations have first members in the category of event or state of affairs.[24]

9.3.2.2 Free Agents Have Active Power

Second, while many naturalists do not employ a powers view of causality, for those who do, it is still the case that the ontology of naturalism knows nothing of active powers. The particulars that populate that ontology are, one and all, exhaustively characterized by passive liabilities with regard to their causal powers. A passive liability is such that, given the proper efficient cause/causal conditions, it is and, indeed, *must* be actualized (*sans* blockers.) As such, the actualization of a passive liability is a passive happening, not an *action*. This fact about passive liabilities is what makes them, along with their causes, fitting entities for subsumption under law. And it is precisely as passive and so subsumable that makes their owners bereft of the sort of first-moving, active spontaneity that is a necessary condition for the exercise of SLA.

Recall our ordering of different sorts of alleged emergent properties in Chapter 2. We showed that the further up the list, the more difficult it is for a faint-hearted naturalist to incorporate the emergent property into his/her overall ontology. This is due to the fact that as one ascends the list, the metaphysical commitments increase and increasingly become excess baggage that is harder and harder to reconcile with the aspects of naturalism discussed in Chapter 2.

Throughout the universe, all natural objects with causal powers possess them as passive liabilities. Again, these liabilities are triggered, manifested, or actualized if something happens to the object and, once triggered, they can produce an effect. For example, dynamite has the power (passive liability) to explode if something is first done to it. And so on for all efficient causes. They are, one and all, passive potentialities. Their actualizations are mere happenings to the relevant object.

But active power is different. In virtue of possessing active power, an agent may *act*, initiate change or motion, perform something, bring about an effect with nothing causing the agent to do so. Active power is not something admitted in the ontology of the hard sciences, period.

9.3.2.3 Free Agents Are First Movers

Third, a "first mover" is a substance that has active power. As such, it is the absolute originator of its actions. Such actions are clear examples of the classic notion of immanent causation. Such an act is not just another caused cause, just one more event in a chain of events in which earlier causes bring about latter effects that, in turn, bring about later effects to form one big series of passive happenings governed by natural law. Such a chain begins outside, runs through, and continues outside the agent with no non-arbitrary, relevant cut-off point to keep the relevant part of the chain "inside" the agent.

No, a first mover is not subject to laws in its initiation of action. Since such an initiation is a first, spontaneous action not caused by a prior event, it amounts to the absolute origination of initiatory movement. Such an origination comes into being instantaneously and spontaneously, and while the effect it produces (e.g., the earliest stages in the raising of one's arm) may well be subject to natural laws, the initiating action is not since there is nothing prior to its coming-to-be on which a law may operate. Moreover, such a first mover is an unmoved mover, i.e., it has the power to bring about an action without having to change first before it can so act. In this way, a first mover provides a natural, *non-arbitrary* regress stopper.

---

[24] Bishop, *Natural agency*, p. 40.

By contrast, since all events in a naturalist ontology are passive happenings, they all are examples of moved movers, i.e., something has to happen to an object first, viz., an event that triggers and actualizes its causal powers, before it can cause something else to happen. In this sense, all naturalistic causation involves changed changers. But a first mover can produce change without having to change first to do so. It should be obvious why such an agent is not an object that can be located in a natural ontology. Unmoved movers with active power are quintessentially unnatural! Indeed, they are exactly like the theistic God in this regard.

It is important to note that the SLA notion of being a first, unmoved mover with active power provides a defeater for a compatibilist strategy (and perhaps certain libertarian strategies like Robert Kane's) long employed to defeat infinite regress problems and to retain various analyses of control. Long ago, compatibilist Nancy Holstrom identified the regress problem in this way:

> If, in order to have control over one's actions, (i.e., for one's actions to be free), it is necessary to have control over the causes of one's actions, i.e., one's desires, then in order to have control over them, it is also necessary to have control over *their* causes, and in order to have control over them, it is necessary … In short, we are lead to an infinite regress, or to a point where the person cannot possibly have any control.[25]

Holstrom goes on to say that this regress runs either backwards in time to events that happened before the agent was born or to events that are outside the agent and, via a causal chain, pass over the agent's boundaries such that the chain continues (spatially) within the agent on to a bodily output of some kind. She offers this solution to the problem: "To go further back [in time or outside the agent] is to leave the person behind."[26] And, says Holstrom, this is problematic because it is the purpose of an account of free will and control to locate the relevant aspects of that account within and only within the agent himself/herself. Thus, such chains are simply beside the point.

Now on the SLA account above, it is obvious why an account of free will must stay within and only within the agent if responsible control is to be properly analyzed. Free actions are paradigm cases of classic immanent causation. But things are not so clear for compatibilists and those libertarians whose accounts employ efficient event-causality and do not include first unmoved movers with active, originative power.[27]

To see this, recall that the causal relations among the members of efficient causal chains are transitive. If an account of control involves such chains, then there is no nonarbitrary place to stop the regress of efficient causes. In such cases, the spatially inside/outside distinction is irrelevant and arbitrary regarding efficient causality and control, though it may be spatially interesting. And the same could be said for diachronic chains that run back in time to events before the agent was born. We believe this same problem besets those who employ passive volitions as necessary causal conditions which, together with one's reason for acting, jointly and efficiently cause that action. And insofar as indeterministic causality is a species of efficient causality bereft of first-moving exercises of active power, we think this problem applies with equal force to chains of indeterministic efficient causes.[28]

---

[25] Holstrom, N. (1977). Firming up soft determinism. *The Personalist*, *58*, 39–51. The quote is from 46–47.

[26] Ibid., p. 47. Brackets ours.

[27] An uncaused volitional advocate has the resources to stop non-arbitrarily the causal chains from "running outside the agent." He/she can stop with an uncaused exercise of active power which by its very nature is uncausable.

9.3.2.4 Free Agents Have Categorical Ability

Fourth, the notion of "categorical ability" in (3) has two important aspects to it. First, it expresses the type of ability possessed by a first mover that can exercise active power and, as such, it contrasts with the conditional ability employed by compatibilists. Second, categorical ability is a dual ability: if one has the ability to exert one's power to do (or will to do) A, then one also has the ability to refrain from exerting one's power to do (or to will to do) A. This means that the circumstances within (e.g., motives, desires, reasons) and outside (environmental conditions) the agent at the time of action are not sufficient to determine that or fix the chances of the action taking place. Given those circumstances, the agent can either exercise or refrain from exercising his/her active power, and this ability is the essential, causal factor for what follows. Among other things, this implies that libertarian acts cannot be subsumed under natural laws, whether construed as deterministic or probabilistic. In this sense they are transcendent and hard to reconcile with staunch or faint-hearted naturalism.

But all the particulars in the naturalist ontology are so subsumable. In fact, all of them are subject to synchronic and diachronic determinism in the sense

> *Diachronic Determinism*: At some time *t*, the physical conditions *C* are sufficient to determine or fix the chances of the next event involving the object and its environment.
>
> *Synchronic Determinism*: At any time *t*, an object's states and movements are determined or have their chances fixed by the micro-physical states of the object and its environment.

This latter determination is bottom/up.

9.3.2.5 Free Agents Act for the Sake of Teleological Ends

Fifth, (4) expresses a view of reasons as irreducible, teleological goals for the sake of which a person acts. In general, we may characterize this by saying that person S F'd (e.g., went to the kitchen) in order to Y (e.g., get coffee or satisfy S's desire for coffee). This characterization of action, according to (4), cannot be reduced to a causal theory of action that utilizes belief/desire event-causation such that reasons amount to efficient causes (or causal conditions) for action. To see this, consider these two sentences:

---

[28] For a discussion of how efficient causal chains relate to the existence of irreducible teleology, see Moreland, J. P. (2002). Miracles, agency and theistic science: A reply to Stephen B. Cowan. *Philosophia Christi*, *4*(1), 139–162.

(a) The water boiled because it was heated.
(b) Smith went to the kitchen because he wanted to get coffee.

(a) is a straightforwardly (efficient) causal assertion. The event cited after "because" (the water's being heated) is the efficient cause for the water's boiling. But while grammatically similar, (b) doesn't employ reasons as causes. This can be seen by paraphrasing (b) as follows:

> (b'): Smith went to the kitchen in order to get coffee. Here, getting coffee is the goal, purpose, end of Smith's free action. Every step he takes (getting out of his chair, walking toward the kitchen, opening the pantry) are means to this end.

If there is anything that naturalists agree upon, it is that there is no such thing as irreducible teleology. Matter is mechanistic, not in the sense that it only engages in action by contact and is bereft of forces, but in that it only behaves according to chains of efficient causes. As philosophers Joshua Hoffman and Gary Rosenkrantz note, attempts to slap teleology onto a naturalist framework really amount to abandonment of naturalism:

> Aristotle's account [of natural function and teleology] does *not* provide a naturalistic reduction of natural function in terms of efficient causation. Nor do characterizations of natural function in terms of an irreducibly emergent purposive principle, or an unanalyzable emergent property associated with the biological phenomenon of life, provide such a reduction. Theistic and vitalistic approaches that try to explicate natural function in terms of the intentions of an intelligent purposive agent or principle are also nonnaturalistic. Another form of nonnaturalism attempts to explicate natural function in terms of nonnatural evaluative attributes such as intrinsic goodness ... We do not accept the anti-reductionist and anti-naturalistic theories about natural function listed above. Without entering into a detailed critique of these ideas, one can see that they either posit immaterial entities whose existence is in doubt, or make it utterly mysterious how it can be true that a part of an organic living thing manifests a natural function ... .[T]he theoretical unity of biology would be better served if the natural functions of the parts of organic life-forms could be given a reductive account completely in terms of nonpurposive or nonfunctional naturalistic processes or conditions.[29]

9.3.2.6 Free Agents Exhibit Top/Down Causation

Finally, libertarian acts involve genuine top/down causation in which a macro-object (a particular person) exercises causal influence in the micro-physical world. As John Searle frankly admits,

---

[29] Hoffman, J., & Rosenkrantz, G. S. (1997). *Substance: Its nature and existence* (pp. 98–99). London: Routledge.

> [W]e are inclined to say that since nature consists of particles and their relations with each other, and since everything can be accounted for in terms of those particles and their relations, there is simply no room for freedom of the will … [Quantum] indeterminism is no evidence that there is or could be some mental energy of human freedom that can move molecules in directions that they were not otherwise going to move. So it really does look as if everything we know about physics forces us to some form of denial of human freedom.[30]

Before we elaborate on Searle's observation, we need to clarify what we mean by *genuine* top-down causation. You may recall that standard meaning of "top/down" causation is actually outside/in causation in which the spatio-temporally wider, structural whole causes things to some of its composing parts. But active power is not a structural property. If there is such, it is an irreducible mental property that emerges from—i.e., is caused by—and is sustained by its subvenient (a widely employed work in philosophy of mind and action theory associated with the lower level base that gives rise to an emergent property) base. If such a property makes a causal contribution to the world, with one exception, it can only be through genuine top/down causation. That one exception is the one we endorse: Active power is exercised by and only by a fundamental, simple soul. As fundamental, substantial souls populate the fundamental level of the mereological hierarchy (if there is such a level) alongside atomic simples.

Returning to Searle's point, consider the following. If the existence of conscious states such as active power is embraced, as it must be if SLA is in view, then a naturalist will have a difficult time avoiding an epiphenomenal depiction of conscious states according to which they are caused by or emerge from the brain, but they are themselves causally impotent.[31] To see why this is so, consider a person getting a drink of water. Now according to the causal closure principle (in tracing the efficient-causal antecedents of any physical event that has a cause, one will never have to leave the physical; if one has to do so, this would be tantamount to admitting a miraculous intervention in the flow of events at the micro-physical level), the cause of the person getting the drink is a relevant brain state.[32] If so, what does the mental state of feeling thirst contribute? If it is a real mental state distinct from the brain state, there appears to be no room for it to affect anything, since the relevant brain state is the adequate cause. More generally, consider the diagram shown in Figure 9.1:

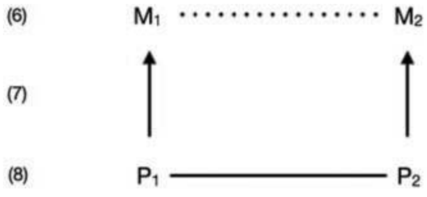


FIGURE 9.1 Mental Causation.

---

[30] Searle, J. (1984). *Minds, brains, and science* (pp. 86–87). Cambridge, MA: Harvard University Press.

[31] See Kim, J. (2005). *Physicalism or something near enough* (pp. 13–22, 32–69). Princeton, NJ: Princeton University Press.

[32] For a critique of non-interventionist accounts of divine (and human) action, along with a defense of an interventionist view, see Larmer, R. (2022). Defending special divine acts. In G. E. Gannsle (Ed.), *Philosophical essays on divine* causation (pp. 174–195). London: Routledge.

Figure 9.1 depicts a sequence of two mental states and two brain states. Let us ask what the cause of M_2 is. If we wish to allow for mental causation, we may say that M_1 is the cause of M_2. However, according to the naturalist principle of supervenience (mental events are determined by and depend upon the relevant physical events upon which they emerge), M_2 supervenes and is dependent upon P_2. Thus, if M_1 is to be the cause of M_2, it will have to cause M_2's subvenient base, P_2. But the naturalist principle of the causal closure of the physical requires that P_1 be the adequate cause of P_2. Moreover, M_1 itself exists in dependence upon its subvenient base, P_1. Thus, assuming the falsity of causal overdetermination (an event has two entirely sufficient causes), we see that there is no room for mental to physical causation (M_1 causing P_2) or mental to mental causation (M_1 causing M_2). The sequence of mental events running through a person's consciousness is like a series of causally impotent shadows.

While they may not be familiar with the argument just given, we have observed above that naturalists Cosmides and Tooby grasp that epiphenomenalism follows from evolutionary naturalism. The best way for a naturalist to avoid the problem of epiphenomenalism is to identify conscious states with brain states, most likely as realizers of types of functional roles. In this way, conscious states can retain causal power because they just are brain states. Unfortunately, this move amounts to a denial of active power in favor of passive liabilities, and with it, a rejection of SLA.

Before we continue our discussion of SLA, we need to take a brief excursus on the impact Kim's argument, especially the role that the causal closure of the physical, plays on genuine mental-event/state causation. Put simple, Kim's argument also rules out the causality of irreducible, intrinsically mental, subjective events/states. This raises two questions that must be kept distinct: (Q1) Should a faint-hearted naturalist accept Kim's argument and the closure principle it entails? (Q2) Does the evidence—in particular, the scientific evidence—support the closure principle? The answer to question (Q1) is yes, and Kim gives the reason why:

> If you reject this principle, you are *ipso facto* rejecting the in-principle completability of physics—that is, the possibility of a complete and comprehensive physical theory of all physical phenomena. For you would be saying that any complete explanatory theory of the physical domain must invoke nonphysical causal agents … It is safe to assume that no serious physicalist could accept such a prospect.[33]

In keeping with our analysis of naturalism as a worldview in Chapter 2, along with its relationship to strong physicalism, Kim is spot on in identifying the real problem here. If genuine mental causation is allowed, then the explanatory power of the naturalist epistemology, along with the Grand Story it certifies, renders it impossible to locate such mental causation in the naturalist ontology. And it stifles an attempt to provide a rationally superior in-principle account (naturalism in either form), relative to competing worldviews of how even all *physical* phenomena obtained an account that is complete and comprehensive.

---

[33] Kim, J. (1998). *Mind in a physical world* (p. 40). Cambridge, MA: MIT Press. Colin McGinn makes the same point. See his *The mysterious flame*. (1999). New York, NY: Bantam Books, p. 93.

Regarding (Q2), a negative answer seems to be what the data support. As Daniel Siegel, world-renowned clinical and research psychiatrist at the UCLA School of Medicine, canonically affirms:

> One of the key practical lessons of modern neuroscience is that the power to direct our attention has within it the power to shape our brain's firing patterns, as well as the power to shape the architecture of the brain itself ... It bears repeating that while the physical property of neurons firing is correlated with subjective experience we call mental activity, no one knows exactly how this actually occurs. But keep this in the front of your mind: Mental activity stimulates brain firing as much as brain firing creates mental activity.[34]

We accept the data affirmed by Siegel. As a result, we take those data to provide defeaters against staunch and faint-hearted naturalism. One may respond that the data only present defeaters for staunch naturalism. But this response fails to assess accurately how inappropriate and *ad hoc* such a faint-hearted adjustment (dualist mental states with causal powers) is. It fails to understand the inner logic of naturalism as expressed in Chapter 2, and it disregards the role that a rival theistic worldview plays in further underscoring the inadmissibility of such an adjustment.

Returning to SLA, it should be obvious why SLA is a feature of the world that a staunch naturalist/strong physicalist must deny. And a faint-hearted naturalist/weak physicalist cannot appropriate SLA into his/her naturalist framework in an epistemically justified or warranted way. There is not nor will there ever be a plausible explanation as to how one can start with dead, brute, non-teleological, law-governed, material particles or stuff with and only with passive liabilities and generate the sort of ontological agent required for SLA by simply re-arranging parts into new external relations.

But maybe SLA is not the best version of free agency. What can we say about this?

9.4 ALTERNATIVES TO SLA

In selecting representatives of alternatives to SLA we have been guided by an attempt to examine important representatives of four different alternatives: Daniel Dennett (*free will irrealism*), John Searle (*creative compatibilism*), Robert Kane (*faint-hearted libertarianism*), and Kevin Timpe and Jonathan Jacobs (*minimalist naturalistic libertarianism*).

9.4.1 Daniel Dennett and Free Will Irrealism

In keeping with Dennett's staunch naturalism and his general approach to consciousness (his Intentional-Stance view), his basic approach to free agency is that, even though all our actions are determined, it is pragmatically useful for society to act *as if* some acts (e.g., child abusers) are done freely and in a morally responsible way. Among the reasons for incarcerating people—rehabilitation, deterrence, protection of society, and punishment—only the first three make sense in a naturalist view of things. Genuine retributive justice is a thing of the past.

---

[34] Siegel, D. (2010). Mindsight: The new science of personal transformation (p. 39). New York, NY: Random House. Cf. Gordon A., with Ziv, A. (2021). The way out: A revolutionary, scientifically proven approach to healing chronic pain. New York, NY: Penguin Random House.

Dennett tries to preserve a remnant of retributive punishment on a staunch naturalist view, but his solution may fairly be taken as a *reductio ad absurdum* against staunch naturalism.[35] Dennett acknowledges that both alcoholics and child abusers are equally determined to act as they do by forces outside their control. If we wished, we could continue to peel back the layers of people engaged in each action and find the genetic or other determining factors producing the behavior. Still, we draw a metaphysically arbitrary line between the two, and we hold child abusers and not alcoholics responsible for their behavior, preferring to treat the latter as a disease.

What justifies us in drawing the line where we do? Dennett explicitly appeals to a utilitarian justification that he takes to be an aspect of our implicit social contract, claiming that it maximizes the greatest amount of good for the greatest number of people (e.g., it serves to deter other acts of child abuse and retain a stable social order constitutive of the contract) if we act *as if* child abusers are responsible but the same cannot be said for alcoholics. Typical of Dennett, we opt for a voluntaristic stance toward certain acts, e.g., child abuse—a big *als ob*—even though libertarian freedom is a figment of an outdated worldview.

Dennett's solution is actually a deflation of the problem, a cure that kills the patient. As such, it is merely a denial of free will combined with the dismissive attitude that, at the end of the day, implies that the irreality of free will is no big deal. We leave it to the reader to decide if Dennett is correct about this.

9.4.2 John Searle and Creative Compatibilism

9.4.2.1 Free Will Within Naturalist Constraints

John Searle proffers a more promising response.[36] Searle represents a serious attempt to solve the problem of free will from within the constraints of a naturalist framework. His view is a compatibilist one, and there are a number of highly respected advocates of this position. We have chosen to examine Searle's account because it is an important, though not widely known position, and it is explicitly formulated in light of neuroscience and a commitment to staunch naturalism (or, perhaps, a version of faint-hearted naturalism very close to staunch naturalism).

According to Searle, for two reasons, we must *posit* a substantial ego as a purely formal, transcendental entity, an entity we *must take* to be real if we are to embrace libertarian action (Searle 33, 56, 73): (i) *reasons-explanations* (e.g., I punched a hole in the paper ballot because I wanted to vote for Bush) cite reasons on or for the sake of which we act (Searle 53, 57), but as such, they are insufficient for explaining why free actions happen. For that we need to postulate a non-Humean (i.e., a non-bundle of perceptions), substantial self purely as a formal, transcendental postulate/precondition for the possibility of free action (Searle 57).

---

[35] Dennett, D. (1984). *Elbow room* (pp. 156–165). Cambridge, MA: MIT Press.

[36] Searle, J. (2007). *Freedom and neurobiology* (pp. 37–78). New York, NY: Columbia University Press, especially 58–78. Page numbers in parentheses in what follows are from this book.

(ii) We *experience ourselves acting freely* in the space of mental gaps in which an earlier mental state is not sufficient for the occurrence of a subsequent mental state (Searle 55–56).

But there are serious problems in embracing the reality of free will itself (Searle 5, 10–11, 17, 38–39, 59, 60). It is hard to conceive how we can harmonize our conception of ourselves as free with the view that the universe consists entirely of mindless, meaningless, unfree, non-rational, brute physical particles. "In the end, perhaps we will have to give up on certain features of our self-conception, such as free will" (Searle 5). Unfortunately, according to Searle, human rationality presupposes free will because rationality must make a difference in our behavior; there must be a difference between rational and irrational behavior, and this can be so only if not all of our actions have antecedent physical causes that are sufficient to determine our actions. Only then can rationality operate, only then is there a certain room for maneuver (Searle 10).

Yet it is hard to see how free will can exist in a world where all events, at least at the macro level, seem to have causally sufficient antecedent conditions. And indeterminacy at the quantum level doesn't help the problem since it is pure chance and randomness which by itself is not sufficient for responsible freedom (Searle 11). Moreover, we are strongly inclined to think that determinism governs the occurrence of natural phenomena, viz., given the right causal antecedents, the effect *had* to occur, and this is incompatible with free will (Searle 38–39). The simple fact is that while we cannot get along without assuming we have (libertarian) free will, it doesn't seem to fit with what we know about the rest of the universe.

At the very least, then, we do not yet know if we have free will and, in fact, it is quite likely that there is no such thing. If, as seems quite likely, the total state of one's brain at t_1 is causally sufficient to determine the brain's total state at t_2, then there is no such thing as free will. So we are faced with this question: Are the causal antecedents of our actions in every case sufficient to determine our actions, or are there some cases where they are not sufficient, and if so, how do we account for those cases (Searle 45)? This appears to be a distinctively philosophical question, but as we shall see shortly, Searle's solution is to dissolve the philosophical problem of free will and replace it with a neurobiological problem.

9.4.2.2 A Solution to the Naturalist Problem of Free Will

Given his skepticism, Searle turns to the task of laying out a solution to the naturalist problem of free will. Searle is clear that mere indeterminacy is not sufficient for genuine free will (Searle 11). Still, if free will is to be real, there must be gaps between antecedent causal conditions and resultant effects in three areas (Searle 42–43): (i) between the reasons for making a decision and the making of it, (ii) between the decision itself and the onset of action, (iii) between the onset of the action and the carrying out of the action to its completion. Phenomenologically speaking, says Searle, we do not experience gaps between conscious states and bodily movements or between physical stimuli and bodily movements. But we do not experience the antecedent conditions as sufficient for the effects in these three areas. Instead, we seem to be aware of gaps in these areas that can be depicted as follows (Searle 59). We actually experience ourselves—whether veridically or not—as acting in these gaps (Searle 55). Figure 9.2 is helpful.[37]

---

[37] Searle, J. *Freedom and neurobiology*, p. 59.

Figure 9.2 should make clear the general contours of Searle's position. After laying out this background, Searle proffers a solution to the problem of free will in a naturalist world. He begins by trying to deflate the severity of the problem with two moves. First, he asserts that we must abandon the old dualist/physicalist dichotomistic language and replace it with a new vocabulary that describes ordinary matter, including the brain's lower-level structure, in the standard ways, and describes consciousness as a typical higher-level, biological, systems property of the brain. By using this new language, we set aside the idea that consciousness cannot naturally emerge when a suitably structured brain obtains, and we treat consciousness as an ordinary case of emergence, e.g., just like the emergence of the properties of digestion.

Second, he deflates the distinctive philosophical problem (e.g., given the nature of brute, mindless, deterministic, mechanistic matter, how is it possible for intrinsically characterized, irreducible consciousness to appear by simply re-arranging brute matter into different spatial configurations?) and replaces it with a neurobiological problem: What exactly are the neuronal processes that cause our conscious experiences, and how exactly are these conscious experiences realized in the brain? (Searle 40, 70). Note carefully, that this question will be answered when detailed correlations are provided between specific brain and mental states, combined with a depiction of indeterminacy necessary for a gap into which free action can be asserted. With this background, Searle tackles the question of how consciousness could move the body (Searle 48), and his answer can be broken down into six steps:

Step one. Given that we have abandoned the old dualist, dichotomistic categories, we can compare consciousness with other higher-level systems properties and find an analogy for how consciousness can move the body. Borrowing an analogy from Roger Sperry, Searle invites us to consider a wheel made entirely of molecules rolling down a hill (Searle 48–49). The solidity of the wheel affects the behavior of each molecule in that its trajectory is affected by the entire solid wheel. Given that the wheel is nothing but molecules, when we say that solidity functions causally in the behavior of the wheel and individual molecules, we are not saying that the solidity is something in addition to the molecules; rather, it is just the condition the molecules are in.

Step two. Searle unpacks the analogy between the wheel and its solidity with the brain and its consciousness: The consciousness of the brain can have effects at the neuronal level

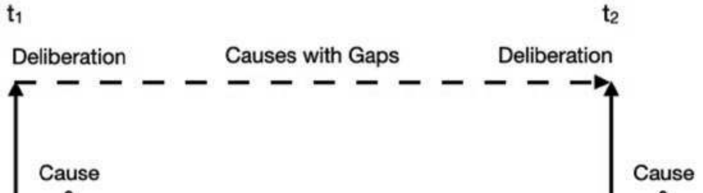


FIGURE 9.2 Naturalist Free Will.

even though there is nothing in the brain except neurons (along with other structured entities such as neurotransmitters). And just as the behavior of the molecules is causally constitutive of solidity, so the behavior of neurons is causally constitutive of consciousness. Thus, when we say that consciousness moves the body, we are claiming that neuronal structures move the body due to the conscious states they are in. Consciousness is a feature of the brain like solidity is of the wheel. To be sure, consciousness, unlike solidity, is not ontologically reducible to physical microstructures, but this is not because consciousness is some extra thing; it is because consciousness has a subjective, first-person ontology and is not reducible to anything with a third-person, objective ontology. But consciousness is causally reducible to neuronal structures in that the former has no causal powers beyond those of the latter. By way of application, one's conscious intention-in-action causes one's arm to rise, but the intention is a feature of the brain system, and at the level of neurons, it is entirely constituted by neuronal activities.

Step three. Searle reminds the reader that reasons explanations (e.g., I raised my hand because I wanted to vote) do not provide sufficient conditions for explaining action, and thus, one must also postulate a self with free will who acts in the gaps listed above (Searle 55). This self is a purely transcendental postulate in Kant's sense, that is, we must act *as if* such a self exists to make sense of free will, but such a formal postulate is not as such a real entity (Searle 58).

Step four. Searle reminds the reader that our problem is a neurobiological one: If free will is real, it must have a neurobiological reality, and there must be some feature of the brain that realizes free will. Consciousness is a system-feature of the brain caused by lower-level entities (neurons, and so forth) and, in turn, realized in the system composed by these lower-level entities (Searle 58–59). Given that free will requires higher-level gaps, how might they be realized at the neurobiological level? Unfortunately, says Searle, there do not appear to be gaps in the brain.

Now, the problem is this: If the total state of the brain at t_1 is causally sufficient to determine the total state of the brain at t_2, then one has no free will. However, if the earlier state is not causally sufficient for the latter, then given certain assumptions about consciousness, one does have free will. How would we build a conscious robot, where at any instant, every facet of its consciousness is entirely determined by states at the microlevel, yet at the same time, the consciousness of the system functions causally in determining diachronically the next state of the system by processes that are not deterministic, but rather, are a matter of free, rational choice? (Searle 70).

Step five. Searle considers two hypotheses as answers to these questions (Searle 61–66). According to Hypothesis 1, the state of the brain at any time is, in fact, causally sufficient for the next state. At the level of conscious deliberation where there are allegedly gaps such that certain antecedent psychological states are insufficient for bringing about the next state (e.g., between deliberating and deciding, between deciding and initiating action, and between initiating and carrying out an action), each such state is matched at the lower level by a sequence of neurobiological events each of which is causally sufficient for the next. Here we have a neurobiological determinism that "matches" a psychological libertarianism. But the psychological level is illusory. There is an experience of free will, but no genuine free will obtains. The experience of such gaps is causally impotent, and, in fact, there are no such gaps. Each lower-level state is sufficient for the next lower-level state and for the upper-level conscious state that synchronically supervenes on it. Free will is an illusory epiphenomenon in Hypothesis 1.

In Hypothesis 2, the absence of causally sufficient conditions at the psychological level is matched by the same absence at the neurological level. The problem, says Searle, is that we seem to have no idea what this could mean. There just are no gaps in the brain. But maybe we need to revise our description of consciousness as a higher-level phenomenon. We must hold on to the idea that consciousness is a feature of the whole system, it is literally present throughout those regions of the brain where it is created by and realized in neuronal activity, and it is located in certain portions of the brain and functions causally relative to those locations. As with the wheel and solidity example, perhaps conscious volitional features of consciousness have effects on the micro-elements of the brain, even though the system is composed entirely of those elements.

So, in Hypothesis 2, (a) the state of the brain at one time is not causally sufficient for its next state; (b) the movement of a state of the brain to the next state can only be explained by consciousness construed as a higher-level system-feature causally but not ontologically reducible to the lower level; and (c) all the features of the conscious "self" are entirely determined at each instant by the lower level. The state of the neurons determines the state of conscious (bottom/up), but any given neuronal state is insufficient for the next such state (lower-level diachronic indeterminism), and the passage of the brain from one state to the next is explained by rational thought processes and volitional choices.

Step six. Currently, says Searle, we have no idea how Hypothesis 2 could be true, but we can set down certain necessary conditions for its being true: (i) Consciousness, as caused by neuronal activity, can cause the body to move. (ii) The brain causes and sustains the existence of consciousness and the "self" (the possession of a unified field of consciousness, the possession of capacities for deliberative rationality, the possession of capacities to initiate and carry out actions) that can make free decisions and carry them out. (iii) The brain is diachronically indeterminate such that the conscious self can make/carry out decisions in psychological gaps with respect to which explanations are provided for the sequence of the brain's states. The indeterminacy at the microlevel is necessary but not sufficient for genuine rationality/choice since the former but not the latter is purely random. How Hypothesis 2 could be true is a mystery, and according to Searle, Hypothesis 1 is more likely to be true (Searle 77).

Because the metaphysics of libertarian freedom (the ontology of SLA noted above, e.g., a first-moving substantial, immaterial self, active power, teleological action, top/down causation that violates the causal closure of the physical and seems to run contrary to bottom/up dependency and determination, a non-empirical faculty of awareness in virtue of which one is able to be aware of one's own self and its invisible exercise of active mental power), almost all naturalists reject libertarian free will. However, a small minority has sought to harmonize such freedom with naturalism, with Searle's account being among the most sophisticated of such attempts. Unfortunately, for several reasons, his attempt must be judged a total failure.

9.4.2.3 A Critique of Searle's Solution

For one thing, Searle's rejection of the dualist dichotomistic vocabulary of mental vs. physical amounts to nothing more than an unwarranted question-begging assertion or, worse, an unmotivated recommendation. Everyone who reads Searle calls him a property dualist. And while he claims to be puzzled by this, he shouldn't be. Searle's descriptions of ordinary matter and the lower-level constitution of the brain are just the ones dualists and physicalists accept. And his description of consciousness is just the one dualists and physicalists accept if it is taken to be a set of *sui generis*, irreducible, intrinsically characterized mental properties. Searle's new vocabulary of ordinary matter and higher-level biologically natural properties is just a verbal shift that expresses in different terms what dualists and physicalists have all along expressed in the "old" vocabulary.[38]

There's nothing new here. There's only one difference in Searle's vocabulary. Most dualists and physicalists use a dichotomistic vocabulary not only to describe two very different sorts of entities, but also to imply that they are so different, that the mental could not emerge from the physical by natural processes. And it is here that Searle's lexicon "justifies" nothing more than an assertion: "Consciousness can emerge from matter." But this begs the question against the majority who eschew such emergence. Searle's depiction of consciousness and ordinary matter is precisely that of the majority, and he provides no argumentation for his assertion besides offering a few bad analogies for consciousness, e.g., solidity, digestion. Alternatively, he may be making a recommendation: "Let's pretend that mind and matter are close enough in their nature that the former can emerge from the latter under suitable conditions." But, again, he does not offer adequate reasons for accepting this dictum, so it is inadequately motivated.

Second, Searle's replacement of the philosophical problem with the neurobiological one is faulty for two reasons. For one thing, the philosophical problem has been around a long time, it is a unique difficulty (How can consciousness come from brute matter? What are we to make of the interaction between two such different orders of being?) in no way identical to the neurobiological one, and it is actually conceptually prior to and presuppositional for the neurobiological issue. After all, if one has good reasons for rejecting either emergence or causal interaction, then the scientific question becomes otiose.

Additionally, Searle's neurobiological question can be answered by listing a set of precise correlations between mental and physical states of the brain. But the explanatory gap is lurking in the neighborhood of such efforts, and far from being a solution to anything, such correlations are precisely what need to be explained. Further, the correlations seem to be radically contingent (e.g., it seems quite possible that zombie worlds and disembodied existence could obtain, and itches as opposed to pains could be correlated with C-fiber firing). Thus, such contingent correlations cry out for explanation, and none will ever be forthcoming within a naturalist perspective. Moreover, such correlations fall far short of being causal laws which require showing that the effect *had* to occur given the cause. Searle accepts this causal necessitation requirement (given the cause, the effect had to occur; see Searle; cf. 69), so he cannot avoid the problem by opting for a weaker analysis of causation.

Third, Searle's employment of Sperry's wheel (steps one and two) as an analogy for top/down causation is deeply flawed. For one thing, the wheel is a whole that is broader than its individual constituent molecules/atoms, but it is not a higher whole whose causal activity on its parts would be a case of top/down causation. Rather, it exhibits outside/in causation at the same level. But consciousness is emergent and *sui generis*. As such, it does ride on top of its subvenient base, so if it exerts causal power, that would, indeed, be a case of top/down causation.

---

[38] For more on this, see Moreland, J. P. (1998). Searle's biological naturalism and the argument from consciousness. *Faith and Philosophy, 15*, 68–91.

For another thing, solidity (and digestion) is a structural property ontologically constituted by subvenient particulars, properties, and relations, so it is easy to see why it is causally constituted by these latter entities—so liquidity is just the subvenient structure into which it is entirely decomposable. But this is not so for consciousness. Given its emergent nature and the transitivity of efficient causality, so-called examples of top/down causation of mental events are both equivocal to the case of solidity and really cases of bottom/bottom causation through mental events.

Regarding the structural property (solidity) and the structural processes of digestion, their causal powers are literally "causally constituted" (Searle's notion) by the behavior of its atomic/molecular parts, and, therefore, are reducible to those lower-level powers. But consciousness is not so causally constituted because it is not analogously ontologically constituted. It is a range of simple emergent properties.

If these properties are to be causally efficacious and we retain the causal closure of fundamental entities at the basic level, in addition to atomic simples (or pieces of gunk), we will have to add simple conscious substances that exemplify the properties of consciousness at the bottom level. Again, if we desire to retain a version of closure, it will need to be transformed from the causal closure of the (micro) physical to the causal closure of fundamental entities at the bottom. We do not accept the closure principle, and, in a latter section, we will address that principle(s) and "top-down" causation. But for now, we have offered two reasons—outside/in vs. top/down causation and structural vs. *sui generis* emergent properties—that show a significant disanalogy between Sperry's wheel, digestion and solidity on the one hand, and consciousness on the other. Consequently, Searle's argument does not go through.[39]

Second, Searle makes a distinction between two types of system features that are emergent. Emergent$_1$ features are caused by micro-level entities and do not exercise independent causality. Emergent$_2$ features are emergent$_1$ and are capable of exercising independent causality once they exist. Searle rejects the existence of emergent$_2$ features because, among other things, they would violate the transitivity of causality. Since he holds that conscious states are emergent$_1$, it is hard to see how those states could have causal efficacy.

Third, as we have already seen, Searle holds to the causal reduction of the mental. In causal reduction, the existence and "powers" of the emergent but causally reduced entity are explained by the causal powers of the reducing, base entities. It is hard to see how he could hold this and avoid epiphenomenalism.

There are also difficulties for Searle's attempt to harmonize his biological naturalism with his treatment of reasons explanations. By explicitly adopting a commonsense approach to such explanations, Searle opens the door for a problem. Recall that he argues, correctly in my view, that properly understood, reasons explanations (Agent S performed act A because of reason R) do not cite causally sufficient conditions in appealing to R. Thus, such explanations are not ordinary causal explanations, and they leave a causal gap which is to be filled by (a transcendental, formally postulated) ego or self who exercises freedom. But this does not explain the role of R in such sentences, and it is evident to common sense—something Searle himself values—that reasons are irreducibly teleological goals/ends for the same of which agents act. Given that we actually do satisfy such reasons explanations, we seem committed to irreducible teleology; indeed, to a form of teleology such that the presence or absence or reasons such as R make a different to the causal history of human persons. It is hard to see how Searle can call his view biological *naturalism* if it countenances irreducible teleology.

---

[39] See Moreland, J. P. (2016). Why top-down causation does not supply adequate support for mental causation. In T. M. Crisp, S. L. Porter, & G. A. Ten Elshoff (Eds.), *Body and soul: Recent debates*, (pp. 51–73). Grand Rapids, MI: Eerdmans.

Finally, Searle's appeal to gaps at the neurobiological level that reflect gaps at the level of consciousness, and which entail that the unfolding states of the brain are not sufficient for the next state does not work. As Kim argued earlier, without such gaps, top/down causation is pre-empted. This is fundamentally due to the ontological dependency of emergent states/events on subvenient ones. The relation of ontological dependency is a transitive one just as is efficient causality. Now, if we grant to presence of the appropriate gaps at the brain level, this doesn't solve anything for two reasons: (1) Given quantum theory, if we accordingly take causal laws to be probabilistic, each state of the brain still fixes the chances of the next state and this would not be the case if top/down causation occurs (in which case, the bottom effect would be guaranteed by the necessitating exercise of active power at the top). So there is a conflict with top/down causation and quantum gaps due to indeterminacy. (2) A conscious state *qua* event is (a) not a continuant, but rather, a fleeting event and (b) entirely caused by the subvenient event with the result that at the level of consciousness, we have a series of mere shadows as far as causality is concerned. Only a substantial continuant can solve the problem, but Searle eschews such if it is taken to be real.

9.4.3 Robert Kane and Faint-Hearted Libertarian Freedom

9.4.3.1 Kane and Naturalistic Constraints

Kane's view is influenced by an attempt to provide a view of libertarian freedom that stays as close as possible to a naturalistic worldview. Evidence for this claim resides in at least three areas of his position: A denial of a substantial self, a failure to identify his central component of voluntary control—exerting volitional effort—with an exercise of active power, and a denial of reasons as teleological goals/ends for the sake of which one acts voluntarily. As we will see, these three areas render Kane's libertarianism problematic and less acceptable than SLA.

The latter two areas will be taken up in due course as we exposit the central contours of Kane's position. Regarding the first area, Kane rejects SD and one of his reasons resides in his eschewal of postulating non-natural entities in developing an account of freedom:

> One crude reaction [to the idea that we are not independent sources of active movements of our physical bodies, and—in defense of our psyches (or ourselves as independent sources of activity)] is to insist that we are not in the natural world at all. The self behind the window is outside the natural world altogether, looking in on it. This reaction is, of course, mind–body dualism of the Cartesian kind. I do not wish to defend such a view here, for I believe, along with most contemporary philosophers, that it is too crude a reaction to the dialectic of selfhood, and creates more puzzles than it solves—but I think it is important to understand why dualism is a natural reaction at a certain stage in the dialectic of selfhood. It is an attempt to salvage the self or psyche as an independent source of activity by placing it beyond the subversive influences of the "world" or "not-self."[40]

---

[40] Kane, R. (1996). *The significance of free will* (p. 94). Oxford: Oxford University Press. Brackets ours.

For Kane, there are two necessary conditions for there to be libertarian free will (LFW):[41]

> *The Principle of Alternative Possibilities* (PAP): A person is morally responsible for what she has done only if she could have done otherwise.
>
> *Incompatibilism*: Determinism is not compatible with PAP.

Given these, the heart of Kane's view is what he calls Ultimate Responsibility.

> *Ultimate Responsibility* (UR): An agent is ultimately responsible for some event *e*'s happening only if (R) the agent is personally responsible for *e*'s happening in a manner that entails that something the agent voluntarily did (or refrained from doing) either caused or causally contributed to *e*'s happening and made a difference as to whether or not *e* took place; and (U) for every *x* and *y* (where *x* and *y* are the occurrence of an event), if the agent is personally responsible for *x* and if *y* is a sufficient condition (e.g., cause, motive) for *x*, then the agent must also be personally responsible for *y*.

UR entails that an agent must be the ultimate originator and sustainer of his/her own actions and goals.

For Kane, the factor that allows for the ultimacy of agent-origination is the occurrence of self-forming actions (SFAs), namely, those occasions of indecision in which people experience a conflict of will, especially in situations involving moral and prudential contexts. SFAs are undetermined, regress-stopping voluntary acts/refraining in an agent's history that are required for there to be UR. A SFA must satisfy three conditions:

> *Self-Forming Action* (SFA): A choice C is a self-forming act *iff* (i) C is made for a reason in that the reason consists, at least partly, in C's being caused by the agent's recognizing that reason; (ii) the agent desires to act more on the recognized reason for C vs. alternative reasons; (iii) whatever C was made, there was at least one alternative possibility.

In order to deal with the problem of luck (roughly, if all the events leading up to the moment of choice still leave the agent with indeterministically open options, then there is nothing in the agent or, indeed, the entire world, that accounts for the difference between choosing one vs. another option; it follows that the choice to do one thing is simply a matter of responsibility-robbing luck), Kane appeals to his notion of an effort of will as an additional requirement to vouchsafe UR. When you choose to do your duty, the choice must result from an indeterministic effort of will that indeterministically causes the choice.

This raises the question of how it is that we are responsibly for an effort of will. Kane answers this question by stating that we are responsible for efforts of will in that they result from character and motives that are themselves indeterministically caused by prior efforts of will, and in that we typically endorse the result of the effort of will. Past and current efforts of will are themselves indeterministically caused and, in turn, indeterministically cause our choices. In this way, we indeterministically cause our SFAs and, thus, are ultimately responsible for them.

---

[41] See Kane, R. (1996); and *A contemporary introduction to free will* (pp. 120–131). (2005). New York, NY: Oxford University Press.

9.4.3.2 Three Objections to Kane's Position

Others have raised several objections against Kane's position, and we do not wish to repeat those here.[42] However, in light of the overall project of this chapter, we raise three difficulties for Kane's view that, if successful, do two things. First, they show that key features of his view render it inferior to SLA. Second, they show that by attempting to develop a faint-hearted version of libertarianism in light of naturalist constraints, Kane is left with a theory bereft of the important factors that make a version of free action a libertarian one.

9.4.3.2.1 Kane's Denial of a Substantial Agent

We have already noted the first problem with Kane's account, namely, his wholesale rejection of an enduring substantial self in favor of either a depiction of the agent as a mereological aggregate or a dynamic system of some sorts. But as we saw, the latter are not genuine continuants nor does the contemporary notion of Immanent Causation, in contradistinction from the classic notion (see, Kane 42–44), solve this problem. It follows that Kane's agent is not capable of (i) being the sustainer of a voluntary action plan for fulfilling one's ends; (ii) being an originator in the only sense that really provides an ontology of the agent apt for grounding authentic LFW, viz., the sort of first moving/unmoved moving, absolute source of action that provides for the needed notion of control.

9.4.3.2.2 Two Problems with Kane's Account of Causality

Second, Kane's account fails regarding causality in two ways. To begin with, his concept of an effort of will is not the same thing as an exercise of irreducible active power. An exercise of active power is intrinsically voluntary and by nature, uncausable by a prior event. This allows for one to be directly aware through introspective attending to the phenomenological difference from such an exercise and a passive event triggered by a prior event. But Kane's efforts of will are not intrinsically voluntary; rather, their voluntariness results from an extrinsic fact that such efforts result in the relevant way from prior indeterministic efforts of will.

Besides the looming problem of choosing between a vicious regress or stopping with a prior effort of will that was deterministically caused by prior states, Kane's account does not give the kind of control needed for LFW as does an exercise of one's active power. His efforts of will are indeterministically caused by prior events rather than being genuinely free exercises of active power that are performed by and, strictly speaking, "up to" the agent himself/herself. Nor are there any introspective, phenomenological data that support the reality of and provide a way of being aware of the crucial occurrence of exercises of active power vs. Kane's associated efforts of will, which are such due to having the right historical pedigree. Kane is aware of this problem and seems to dismiss it as unimportant:

---

[42] See Goetz, S. (2008). *Freedom, teleology and evil* (pp. 19, 22, 29–31, 53–55, 64–66). New York, NY: Continuum; Clarke, R. (2017). Incompatibilist (nondeterministic) theories of free will. *Stanford Encyclopedia of Philosophy*.

> Another objection to the preceding theory [Kane's theory of free will] is that we are not consciously aware of making competing efforts [of will] when we engage in self-forming choices. But the theory does not require that we be consciously aware of these competing efforts. The idea was to compare exercises of free will to other cases of parallel processing in the brain, such as vision. Neuroscientists tell us that when we see a visual object, such as a red barn, the brain actually processes different properties of the object (like shape and color) separately, through parallel pathways whose results are eventually brought together in the visual image. We are not introspectively aware of processing the redness of the barn and its shape separately and in parallel … But if these neurological theories are correct, that is what we are doing.
>
> The preceding account of free will is suggesting that something similar may be going on when we make self-forming choices. We are not introspectively aware that our efforts (our efforts to make one or another of our competing choices succeed) are being processed on separate, though interacting, pathways in the brain; but that process may in fact be what is going on.[43]

In our view, it is often legitimate to postulate something to explain a set of facts that stand in need of such an explanation. But when it comes to postulating things of which one should have introspective awareness, especially something like engaging in competing efforts of will resulting in an SFA, it seems like one should have such an awareness, both of the struggle of competing desires, values, and alternatives, and the effort taken to resolve the conflict. This would seem to be a prime case where introspective awareness was available and essential to knowing that the choice was a free one.

A second problem with Kane's causal account is his rejection of a causal-powers ontology in favor of a causal view that permits indeterministic causation. We recognize that there are several views about causality being debated in the current literature. But it seems to us that an ontology of causal powers is a necessary condition for any version of LFW worthy of the name. Why? If an agent has causal powers, among them active powers, then an agent-exercise of such provides the only, or at least, the best way to establish indeterministic control of the agent's actions. Such exercises are, indeed, genuine, non-arbitrary regress stoppers.

There are two characteristics of an exercise of active power that are crucial here. First, by their very nature, they cannot be caused by prior events. If they are caused at all, it is a substantial agent acting as a first, unmoved mover who does so. With apologies to the early Chisholm, this solves the regress problem. Second, *sans* overriding blockers, such an exercise produces, brings about, necessitates its effect(s), and that is another reason why an agent who exercises active power has the sort of control needed to be a responsible free agent.

It is clear that indeterminate causation lacks this necessitating factor. Consequently, Kane's view suffers from the problem of luck in a way that the SLA account does not. Given worlds W_1 and W_2 exactly alike in all respects up to the moment of choice A, the agent does A in W_1 and refrains from doing A in W_2. There are no factors in Kane's account that solve this problem.

---

[43] Kane, *A contemporary introduction to free will,* p. 144. Brackets ours.

The advocate of SLA suffers from no such problem. On this view, the difference between doing A in W_1 and refraining from doing A in W_2 is due to how the agent exercised his/her free will. One determines oneself which alternative to pursue, and in doing so, the agent is the ultimate source and first moving originator of that alternative. In this way, it is literally “up to the agent.” If someone asks why the agent did A instead of refraining to do A, it may be that the reason for A was more important than the one for refraining, but at some point, the only answer needed is that the agent just chose a certain alternative and that’s it. There is no difference making further fact of the matter.

If one responds that this makes the choice an arbitrary one and, consequently, undercuts the agent’s responsibility for the choice, the SLA advocate will reply that, quite the contrary, this is exactly what makes the agent responsible. The agent just acted in the relevant way of being in control of that act, and any further causally relevant fact of the matter—e.g., the need to cite some reason R as a necessary causal condition—would rob the agent of being the sole, ultimate, responsible, controlling source of the act.[44]

Moreover, just like it makes no sense to ask what caused God on the assumption that God is a necessary being, so it makes no sense to ask what factor caused the SLA agent to act since that act is a first-moving one of an unmoved mover. In both cases, it is appropriate to say that the buck stops there and that’s the end of the matter.

9.4.3.2.3 *Kane’s Rejection of Reasons as Teleological Ends*

The third problem is that Kane’s view follows a typical pattern by explicitly rejecting reasons as teleological ends; rather, they are efficient causal conditions. Besides the fact that Folk Ontology and introspective awareness of the teleological nature of reasons for the sake of which we act, a construal of reasons as teleological ends provides an excellent explanation for what it is that guides the ordering of an action plan, and adjustments to that plan in light of things that require taking an alternative pathway, by being that stable end for the sake of which the plan is ordered and adjusted in end-appropriate ways. However, it is hard to see how a chain of mere efficient causes or efficient causal conditions can serve as an adequate explanatory substitute or this teleological view of the role of reasons as action guides.

9.4.4 Kevin Timpe and Jonathan Jacobs: Minimalist Naturalism and Libertarianism

Dennett and Searle are representative of most philosophers in rejecting the plausibility of combining naturalism and libertarian free agency. Very few philosophers think that it may be possible to present a minimalist version of both naturalism and libertarian free will that some naturalists would accept. However, Timpe and Jacobs believe they have done exactly that.[45] We remain skeptical because we think that any such attempt will (1) be so minimalist about both items that it will be questionable as to whether the minimalist view is a genuine representation of naturalism and libertarianism; (2) involve advocates who help themselves to entities (e.g., some mentioned in our description of SLA above) that a savvy naturalist, aware of the epistemic and Grand-Story location constraints on a naturalist ontology, would eschew. Too often, Timpe and Jacobs simply embrace such entities with mere assertions that they aren't problematic.

---

[44] For an excellent discussion of these points in the context of analyzing Reid’s view of LFW and in light of a non-causal libertarianism that focuses on the nature of an exercise of active power, see Rowe, W. (1991). *Thomas reid on freedom and morality* (pp. 162–184). Ithaca, NY: Cornell University Press.

[45] See Timpe and Jacobs.

For example, Timpe and Jacobs claim that free will is an emergent systems-level feature.[46] But as we saw in Chapters 2 and 7, an appeal to systems and system-level emergent properties is problematic for a number of reasons, especially in light of the presence of a number of staunch hylomorphic views that entail several ontological commitments that seem to be defeaters for any positive version of naturalism (one that does not merely deny God's existence) worthy of the name.

However, even if such a minimalist account does show that libertarian free will is "consistent with" or "not incompatible with" naturalism or that naturalism "doesn't rule out" libertarianism, we're not sure this amounts to much worth having. The proper question is not, "Can a minimalist version of naturalism and libertarianism be shown to be logically consistent?" Rather, it is "Given the most reasonable form of naturalism and theism as rival worldviews, is it more reasonable than not to believe that the existence of libertarian actions and agents are more at home in a naturalist worldview than a theistic worldview? What is the truth of the matter in the actual world?"

As Timpe and Jacobs admit, there is a near consensus that naturalism is "certainly incompatible" with libertarian free will. Maybe that consensus is not due to people being unfamiliar with the literature on action theory. Perhaps people rightly grasp the inner logic of what naturalism ought to be if it is to claim superior epistemic credentials and explanatory power. And perhaps through introspective knowledge by acquaintance of active power and the teleological role of reasons/motives in life, people have a clear concept of what libertarian freedom really amounts to even if, for ideological reasons, they do not believe in it. We have tried to show that this is the case. Whether or not we have succeeded, we believe the questions we propose are the right ones to ask.

In this context, it is important to recall a point made by Daniel Stoljar noted in Chapter 2: Those who do not accept the idea that staunch naturalism (our words) virtually requires acceptance of strict physicalism do not make a conceptual mistake—e.g., some minimalist version of libertarianism may well be conceptually consistent with faint-hearted naturalism (our words). Instead, according to staunch naturalists, they are flying in the face of a factual thesis (warranted by science) we have overwhelming reason to believe.[47]

In sum, it may be possible to present a minimalist version of both naturalism and libertarian free will that a few naturalists would accept. But these versions may be irrational in light of factual considerations and, moreover, they may be so watered down that they distort naturalism and libertarianism beyond what is acceptable.

---

[46] Timpe and Jacobs, p. 330.

[47] Daniel Stoljar, *Physicalism* (London: Routledge, 2010), chapter one.

9.5 CAUSAL CLOSURE OF THE PHYSICAL AND “TOP-DOWN” CAUSATION

9.5.1 Kim’s Supervenience Argument

The positive argument against top-down causation we have in mind is one that appropriates certain insights from a general case against such causation by Jaegwon Kim.[48] Kim’s supervenience argument (a.k.a. the exclusion argument) purportedly shows that, given the irreducibility of the mental, there can be no mental causation in a world that is fundamentally physical as depicted in the mereological hierarchy, and according to Kim, this raises serious problems regarding cognition and agency, two features of our lives that are hard to give up. The supervenience argument, says Kim, may be construed to show that mental causation is inconsistent with the conjunction of four theses: (1) closure; (2) exclusion (no overdetermination); (3) supervenience; (4) mental irreducibility. Kim invites us to consider two physical events, *p* and *p**, along with two mental events, *m* and *m** such that (1) *m* and *m** supervene on *p* and *p**, respectively (where supervenience includes the notion of dependence and determination, even if this is not taken to be efficient causality) and (2) *p* causes *p**.

The argument proceeds in two stages. *Stage 1*: Focus on *m* to *m** causation. Since *m** obtains in virtue of *p**, if *m* causes *m** it must do so by causing *p**. *Stage 2*: Kim offers two different ways to complete the argument, the first of which is most relevant for our purposes. Assuming causal closure and exclusion (no causal overdetermination), *p* will be the cause of *p** and there is no room for *m* to be involved in bringing about *p**. We have *m* and *m** supervening on *p* and *p**, respectively, and *p* causing *p**, nothing more and nothing less.

9.5.2 Kim’s Crucial Background Assumption

This argument has been widely discussed, but what have garnered less attention are certain background issues Kim advances prior to formulating the supervenience argument. According to Kim, the fundamental idea behind the argument is what he calls

> *(Jonathan) Edwards’s Dictum*: Vertical determination excludes horizontal causation.[49]

In support of *Edwards’s Dictum*, Kim invites us to consider a lump of bronze whose macro-properties such as color, density, or shape are dependent on and determined by its synchronous micro-structure. Now, suppose the lump is yellow at time *t*. Why is this the case? Two answers are possible: (1) Its micro-structure at *t* causes the color to emerge. (2) It was yellow at $t–\Delta t$. Unfortunately, says Kim, there is a tension between the two such that there is neither need nor room for the second explanation. We would add that this seems especially correct, given a causal-powers view of causation. The lump at *t* will be yellow, given its micro-structure, no matter what was the case at $t–\Delta t$.

---

[48] Kim, *Physicalism or something near enough*, pp. 32–69. Cf. Buckareff, A. (2012). An action theoretic problem for intralevel mental causation. *Philosophical Issues*, *22*, 89–105. For a number of reasons, we do not accept the causal closure of the physical. In this regard, see Lowe, *Personal agency*, pp. 58–78; Swinburne, R. (2013). *Mind, brain and free will* (pp. 104–127). Oxford: Oxford University Press; Hasker, W. (1999). *The emergent self* (pp. 58–80). Ithaca, NY: Cornell University Press. Our point here is that causal closure is something a staunch naturalist should accept.

[49] Kim (2005, p. 36).

Kim advances a second illustration taken from Edwards. The successive images of an object in a glass do not remain the same nor do earlier images cause latter ones. Rather, each successive image is caused by, and the series of images is constantly renewed with respect to, a sequence of new light rays that bring about each image. Again, there is neither need nor room for each image to cause its successor or to exert top-down causation.

In the arguments to follow, we will be employing Edwards's Dictum and the two cases used to illustrate it in my arguments against top-down emergent-property causation, and we will be considering situations in which a mental property emergently supervenes on a subvenient physical base. Some might think that these supervenient situations are not analogous with the second case (the glass-image case) above. Why? Of the two cases, the latter one (the glass-image example) is a case of efficient causality and the former one (the bronze-color case) is a case of synchronous supervenience. And, the objection continues, efficient causality is just different from emergent supervenience, so the glass-image case is not relevantly analogous with cases of emergent supervenience.

Some take synchronous supervenience to be an example of efficient causality, but even if we don't agree with this opinion, the glass-image case is still relevantly analogous with the emergent supervenient situations because both sorts of cases (efficient causality and the relation between cause and effect; emergent supervenience and the relation between emergent and subvenient properties) involve an ontological dependence/determination of one entity on another (diachronically, synchronically) and we believe it will become evident below that this is the important analogous feature between the two cases for my purposes.

Note that *Edwards's Dictum* and the two supporting illustrations are not defeated by the truth of a causal-powers view of causation and properties. For example, even if the instantiation of secondary qualities brings along their individuating causal powers, those powers are preempted and never actualized. Moreover, given a rejection of overdetermination, it is important to see that in both of these cases, top-down causation is eliminated without a fundamental commitment to the closure of the physical, a commitment we do not embrace. It is emergence as depicted in the mereological hierarchy that does the trick, along with the fact that we have a pretty good micro-physical story depicting sequences of change at that level as a causal chain. Instantiations of properties at the emergent level are completely dependent on and determined by instantiations of properties at the micro-level; changes at the emergent level are due to changes at the micro-level. Given emergent dependence, the emergent level becomes a series of epiphenomenal states synchronously dependent on the micro-level. Diachronic causation at the emergent level is preempted. And given that there is a causal chain at the micro-level, there is neither need nor room for top-down causation.

9.5.3 The Major Difficulty: Emergence

In light of Kim's background considerations just mentioned, it would seem that the culprit regarding top-down causation is emergence. Granting the mereological hierarchy's twin commitments to the micro-physical level being basic and to genuinely emergent properties, these properties or, rather, their instantiations, seem to be a series of epiphenomenal states. This surely appears to be the case with secondary qualities and value properties, and it would be odd, to say the least, to make an exception in the case of mental properties. Even if the instantiation of these three sorts of properties brings along their causal powers, there doesn't seem to be a need or room for their exercise.

9.5.3.1 No Causal Overdetermination

In the remarks just offered and those to follow, we are presupposing that there is no such thing as

> *Causal Overdetermination (COD)*: An event *e* is causally overdetermined *iff e* is wholly or sufficiently caused by both *X* and *Y*,

especially in cases in which one cause (a baseball) is ontologically dependent on the other cause (a collection of micro-physical parts composing, or allegedly composing, the ball). But COD has been defended by Theodore Sider in the context of discussing Trenton Merricks's eliminativism regarding nonliving macro-objects.[50] Now it seems to us that many, perhaps most, philosophers reject COD. Accordingly, we are tempted to rest content with limiting the effectiveness of my arguments to those who agree with me in this regard. But we believe there is more to be said here, and we want to raise two difficulties with the notion of COD.

To understand the first difficulty, consider the following statements from Helen Steward: "Supposing that substance causation does indeed make sense, there is a question whether substance causation by a big and complex substance like an animal could ever fail to be constituted by causation going on at the level of the much smaller and simpler substances that are the animal's *parts* ... that is to say, whether we can really understand what might be called *top-down* causation, where causality by the complex does not merely reduce to the sum total of a lot of causality by the simple."[51] Elsewhere she states:

> The challenge, as I see it, is to understand how on earth it can be that the animal has any real, independent efficacy *of its own*: an efficacy that does not merely reduce to the efficacy of its various parts.[52]

Steward's worry can be expressed in the following principle:

> *Complex Whole Causation* (CWC): The functioning of some complex object O composed of separable parts e_1, e_2, ..., e_n to produce some effect E consists in the coordinated functioning of e_1, e_2, ..., e_n (or some relevant subset) each of which produces a sub-effect of E, and E just is the sum of such sub-effects.

Given this principle, it seems that the causal work in, say, breaking a window is accomplished by its atomic constituents, and there is no work for the ball as a whole to do and no room for that work. Indeed, when we say that the ball broke the glass, we would more accurately say that the ball broke the glass in virtue of its atomic constituents, and this second assertion seems straightforwardly to give way to a paraphrase along the lines of CWC in which the atomic constituents themselves are identified as the true cause without remainder. And if we add to CWC an epistemic principle of simplicity, then given that we have a clear story about the causal role of the baseball's constituents in breaking the glass, we have no epistemic justification for multiplying causes and postulating the baseball as a whole as an additional cause.

---

[50] Sider, T. (2003). What's so bad about overdetermination? *Philosophy and Phenomenological Research* 27: 719–726.

[51] Steward, H. (2012). *A metaphysics for freedom* (p. 206). Oxford: Oxford University Press.

[52] Ibid., p. 227.

9.5.3.1.1 Ted Sider's Counterargument and Our Response

Sider responds by saying that on several views of causation, the baseball and its atomic constituents qualify as causes: counterfactual, covering law, probability-raising, and primitive-causal-relation analyses. Now while this may be true, we are assuming a causal-powers analysis of causation, and Sider does not explicitly mention it. And a causal-powers view seems to support a rejection of COD because if, say, the atomic constituents of a ball actually bring about an effect, there just does not seem to be need or room for the ball as a whole to do anything.

This claim can, perhaps, be clarified by considering a view of causation that Sider mentions in order to set it aside. On this view, causation is a kind of fluid divided among the various potential causes of an effect. If one such cause produces the effect, the fluid is used up and no other potential cause can act. Atoms that break a glass use up the available fluid, leaving none for the baseball itself to use.

Now while the fluid example is too crude a picture of causation to be accepted, it does bear an important analogy with a causal-powers position. Just as the fluid in Sider's example brings about or produces its effect, so causes on a causal-powers view bring about or produce their effect. In the fluid case, the problem is not that there would be no fluid left over for another purported cause to use. Rather, the problem is that if one stream of fluid is adequate to produce an effect, say water from a hose knocking a quarter off of a fence, there would not be room or need for a second stream of fluid to do anything, especially if that second stream was somehow ontologically dependent on the first one.

In addition, Sider does take up the epistemic problem of simplicity, but he does so in the context of someone wielding the principle to deny the existence of macro-objects like balls. His response is that the epistemic principle is not an argument against the existence of macro-objects—after all, we may have other reasons besides causality to believe in their existence—rather, it is an argument against one argument for them, viz., that they must exist because they exhibit causal activity in their own right. But we are not using CWC to argue against the existence of macro-objects. We are arguing that such wholes do not have causal powers in addition to those of their atomic parts such that COD is exhibited. And Sider seems to agree that the epistemic argument does, in fact, count to some degree against COD.

9.5.3.1.2 Macro-Objects Are Aggregates

Our second argument against COD is this: Granting (as we do) the existence of macro-objects such as a baseball, the metaphysical constituents added to their atomic simples to give such macro-objects being are not the sorts of entities that ground additional causal powers for such objects besides those that follow from their atomic simples. The sorts of macro-objects of interest are mereological aggregates, which we defined in Chapter 3 as follows.

> *Aggregate*: An object *O* is an aggregate *iff* (i) *O* has at least two separable parts, and (ii) the separable parts of *O* stand only in external relations between each other and *O*.

An aggregate is a particular whole that is constituted by (at least) separable parts and external relation-instances between and among those separable parts.[53]

Now a baseball breaks a glass due to the additive causal powers of its atomic simples. The relational structure (and its surface) is not the sorts of entity that adds to the ball's causal powers. It could be argued that this point counts against alleged cases of COD that involve entities like baseballs because they just are a collection of parts standing in a structure. But when it comes to the instantiation of genuine emergent properties, such as mental properties, these mental states are not structural wholes composed of atomic simples, and these states, or more precisely, their properties, bring along with them their own causal powers when instantiated, and such powers could, at least in principle, be exercised, maybe even in ways that satisfy COD. It is to this issue that we now turn.

9.5.3.1.3 Adapting a Defeater from O'Connor and Churchill

O'Connor and Churchill advance the following Kim-style argument against non-reductive physicalism that is important for our discussion:

(1) *Causal non-reductionism*: causation is a real, irreducible relation.
(2) *Production account of causation*: causation consists in the exercise of causal powers.
(3) *Causal theory of properties*: properties are individuated with respect to causal powers.

(1)–(3) unpack a causal powers ontology according to which causality is irreducibly real, it consists in the production of an effect, and causal powers are part of the identity conditions for properties.

(4) *Supervenience thesis*: mental properties strongly supervene on physical properties.
(5) *Realization thesis*: mental properties are realized by physical properties.
(6) *Completeness/causal closure of physics*: every physical event that has a cause has a complete physical cause.

(4)–(6) unpack nonreductive physicalism in which mental properties are either functionally or structurally supervenient. (1)–(6) are inconsistent with (7):

(7) A mental property, M, is distinct from its physical realizer property (or properties), P, and each event that consists of M's being instantiated exercises a distinctive form of causality that in one way or another impinges on the realm of physical events (assumption for reductio).
(8) The instance of M either (a) directly produces a subsequent mental event, M*, or (b) it directly produces a wholly physical event, P*.

(5) The *realization thesis* and (2) *causal production* seems to rule out (8a) because they jointly imply that M has to cause P* on which M* is ontologically dependent; M cannot directly cause M*. So,

(9) Not (8a).

---

[53] There is a debate as to whether or not one should add an additional constituent, viz., a surface or boundary to the analysis. For our purposes we needn't take a stand on this issue.

(6) *Completeness/causal closure of physics*: every physical event that has a cause has a complete physical cause.

(10) If (8b), then the physical event P* is overdetermined by M and some other physical event.

Now, P* could be counterfactually dependent on M&P, but since we are assuming a causal production view of causation, there is no need or room for M in light of P. So we get

(11) There is not systematic mental-physical overdetermination, as the consequent of (10) implies.

(11) and the consequent of (10) imply not (8b) by *modus tollens*.

Therefore,

(12) M does not make a distinctive contribution to occurrences in the physical world, whether wholly physical or supervening mental occurrences.

O'Connor and Churchill conclude by claiming that since (3) *the causal theory of properties* rules out epiphenomenalism, we should either identify M with P or just deny that M is a genuine property.

We believe this argument is successful against nonreductive physicalism. But our purpose in presenting the argument is not to evaluate it on that score. Rather, we want to adjust the argument to show that the adjusted version is successful against the causal efficacy of emergent property dualism. Let's begin with a few observations. First, as far as we can tell, the argument does not make or, at least, require explicit use of the supervenience thesis (4). In (12), the phrase "supervening mental occurrences" could just as well have been "realizable mental occurrences." And the argument's move from (8) to (9) employs realization, not supervenience. In our view, realization is sufficient to go from (8) to (9), but it is not necessary. Our investigation of Edwards's dictum above suggested that supervenience, especially emergent supervenience, alone, without the need to appeal to closure, will justify the move from (8) to (9). The problem with the causal efficacy of an instantiation of an emergent property is that it is a state of affairs totally dependent on and determined by its subvenient state.

But what about O'Connor and Churchill's claim that (3) the causal theory of properties rules out epiphenomenalism? If this were the case, it would count against (3). Why? Because we have strong intuitions that there is a relevant possible world in which epiphenomenalism is true, say, a world in which a mental property's causal powers are not, and perhaps, cannot be exercised, even though they are present. In such a world, those powers are real but preempted. Remember, it is not enough to avoid epiphenomenalism that an instantiated emergent property has causal powers. Those powers must be capable of being actualized, and if for some reason or another—such as Edward's dictum and the two supporting illustrations cited above—that cannot happen, we have both (3) and epiphenomenalism. Thus, we think their premise (7) should be replaced with (7′) (we have italicized our changes to (7)):

> (7′) A mental property, M, is distinct from its physical *subvenient* property (or properties), P, and each event that consists of M's being instantiated *has the per se capacity to exercise* a distinctive form of causality that one way or another impinges on the realm of physical events.

According to (7′), a causal theory of properties means that the instantiation of a mental property M carries with it the capacity for M’s unique causal powers to be exercised, a capacity that can be blocked or preempted by other factors. In light of these observations, we believe a causal-powers Kim-style argument can be developed against the causal efficacy of emergent mental properties along the lines of the O’Connor/Churchill argument. This argument accepts premises (1)–(4) (with a slight adjustment to (4)), sets aside (5) and (6) and uses (7′) instead of (7). It also involves stating a new premise NP for the purposes of reductio. Here is the argument:

(1)–(3) as before (a causal-powers ontology).

(4′) The conjunction of the supervenience thesis and the mereological hierarchy’s depiction of the micro-physical level as basic and its sequence of events as a causal sequence.

(7′) A mental property, M, is distinct from its physical *subvenient* property (or properties), P, and each event that consists of M’s being instantiated *has the per se capacity to exercise* a distinctive form of causality that one way or another impinge the realm of physical events.

(NP) An instance of M makes a distinctive contribution to occurrences in the physical world, whether wholly physical or supervening mental occurrences.

(8) An instance of M either (a) directly produces a subsequent mental event, M*, or (b) it directly produces a wholly physical event, P*.

(4′) (the supervenience and mereological-hierarchy thesis) and (2) (causal production) seem to rule out (8a).[54] So,

(9) Not (8a).

(10) If (8b), then the physical event P* is overdetermined by M and some other physical event.

---

[54] Another way to see why emergent supervenience rules out M to M* causation is to consider the fact that on a causal-powers view, a cause produces an effect in virtue of the actualization of the cause’s dispositional powers. Now, such an actualization will require the disposition to pass from potency to actuality through some time frame of t_1 to t_2. From t_1 to t_2, the dispositional power passes through various degrees of actualization until it reaches full actuality at t_2. Suppose at t_1, M has the dispositional potential D_p to bring about M*, and at t_2, M has D_a, the full actualization of D_p. What is happening at the subvenient physical level during the interval from t_1 to t_2? It seems like two scenarios are possible. *Scenario one*: The subvenient state remains the same. But if this is true, supervenience is violated because the subvenient state at t_1 underlying M&D_p is the same as that underlying M&D_a at t_2. In this case, there is a mental difference without a subvenient physical difference, and this physical base is not sufficient to fix the supervenient state. *Scenario two*: At each degreed state between t_1 and t_2 in which D_p is increasingly becoming more and more actualized until D_a is reached, there is a different subvenient state. However, in this scenario, the transition between D_p and D_a is not the actualization of a genuine causal power. Rather, it is just a series of gradually different epiphenomenal states fully dependent on/determined by the graded differences at the subvenient level. Thus, supervenience rules out M to M* causation because in scenario one, supervenience would be violated. And on scenario two, the transition between M&D_p to M&D_a is like a series of shadows. All the ontological work is being done at the subvenient level.

Assuming a causal production view of causation and (4′), there is no need or room for M's causal activity in light of P. So we get

> (11) There is not systematic mental-physical overdetermination, as the consequence of (10) implies.

So we conclude that not (8b) and, more generally,

> (12) An instance of M does not make a distinctive contribution to occurrences in the physical world, whether wholly physical or supervening mental occurrences.

Notice that (12) is inconsistent with (NP) and, thus, (12) completes our reductio. Epiphenomenalism follows for all possible worlds in which mental properties satisfy (4′). This is consistent with (1)–(3). In these possible worlds, mental properties, as it were, bring their defining mental powers along with them. It's just that emergent supervenience and the mereological hierarchy preempt their actualization.[55]

---

[55] Cynthia and Graham Macdonald attempt to provide an account of a specific form of downward causation for emergent properties (specifically, the instances of such properties) in the special sciences, including mental properties, that preserves the causal closure of the physical. See Cynthia Macdonald and Graham Macdonald, "Emergence and Downward Causation," *Emergence in mind*, ed. by Cynthia Macdonald and Graham Macdonald (Oxford: Oxford University Press, 2010), 139–68. Working within a framework of immanent universals and a Lombardian account of event existence and identity, of causality (x and y stand in the causal relation to each other irrespective of the descriptions under which they fall). (2) Just as the exemplifying of a determinable (being a color) and one of its determinates (being red) constitute a single event that is one exemplifying of two properties, so there is a single event consisting in the exemplifying of a supervenient (e.g., mental) property and a subvenient (e.g., physical) property (the Co-Instantiation Thesis—two or more properties of an event can be co-instantiated in a single instance). (3) By extensionality, if the physical event (the instantiating of the subvenient physical property) is causally relevant/efficacious, so is the mental one (since they are the same event). (4) If a property has instances that are causally efficacious, then the property has causal powers, so mental properties have causal powers. (5) In this way, the causal relevance of mental properties is preserved without violating physical causal closure. Here is our brief response to the Macdonalds: First, the instantiation of being a color at a time is not identical to the event of the instantiation of being red at that time. Rather the former is an essential constituent of the latter. Second, the determinable/determinate relation is a bad analogy for the emergent supervenient property/subvenient property relation. The former is a constituent/whole relation, and the latter is a dependency relation between two different properties and their instantiations. Even if the instantiation of a case of the former is regarded as one event, an instantiation of the latter is two distinct events standing in a dependency relation. Some (e.g., John Searle) take the relationship between an emergent-property event and a subvenient-property event to be causal. Even if this is wrong, it is not unintelligible. If the emergent-property event was identical to the subvenient-property event, then the causal interpretation would, in fact, be unintelligible since a causal event is not identical to an effect event. And given a causal-powers ontology, emergence, the hierarchy and closure, the physical event and not the mental event is causally efficacious. Finally, if one insists on taking the instantiation of the two properties to be a single event, it still seems that this event is efficient-causally efficacious *qua* being physical and not *qua* being mental, even if the mental property tracks (e.g., due to multiple-realizability) a different causal profile than the former. In light of a causal powers approach, it is the causal efficacy, not the causal relevance, of the mental that is required to preserve mental causation and responsible agency. For a brief discussion of the difference between causal efficacy and relevance, see Ravenscroft, I. (2005). *Philosophy of mind: A beginner's guide* (pp. 144–155). Oxford: Oxford University Press.

9.6 CONCLUSION

Jaegwon Kim observes that,

> if a whole system of phenomena that are *prima facie* not among basic physical phenomena resists physical explanation, and especially if we don't even know where or how to begin, it would be time to reexamine one's physicalist commitments.[56]

We have argued that genuinely non-physical mental entities, including the ontological commitments entailed by SLA, are the paradigm case of such a system of phenomena. And if we have succeeded in showing that SLA is the preferred view of free will, it follows that we are simple, substantial, enduring souls.

---

[56] Kim, J. (1998). *Mind in a physical world* (p. 96). Cambridge, MA: MIT Press.

PART V

NEW AND NEGLECTED RESPONSES TO COMMON DEFEATERS AGAINST SUBSTANCE DUALISM

CHAPTER 10

Important Frequently Raised Defeaters Against Substance Dualism

We have presented our case for Substance Dualism (SD) and as a part of that presentation, we have responded to defeaters raised against SD and provided defeaters for alternative views. However, there is a set of regularly proffered objections to SD that we have postponed addressing until this penultimate chapter. Accordingly, while we must be selective, we will try to respond adequately to the most important problems often raised against SD.

Before doing so, it is important to recognize that some who don't ascribe to SD are now arguing that standard anti-SD arguments are not as triumphant as previously thought. After working through standard arguments against SD, José Gusmão Rodrigues, himself a Russellian monist, concludes:

> A significant part of current literature on the mind-body problem, personal identity and other general metaphysical questions seems to me rife with anti-dualistic assumptions without good arguments to back it up. Such an attitude is dogmatic and contrary to good philosophical spirit.[1]

Likewise, William Lycan observes,

> Being a philosopher, of course I would like to think that my stance is rational, held not just instinctively and scientistically and in the main-stream but because the arguments do indeed favor materialism over dualism. But I do not think that,

---

[1] Rodrigues, J. G. (2014). There are no good objections to substance dualism. *Philosophy*, *89*(2), 221–222.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

> though I used to. My position may be rational, broadly speaking, but not because the arguments favor it: Though the arguments for dualism do (indeed) fail, so do the arguments for materialism. And the standard objections to dualism are not very convincing; if one manages to be a dualist in the first place, one should not be much impressed by them.[2]

Last, physicalist Karen Bennett is painfully honest where she writes, “Unfortunately, it [physicalism] is closer to an article of faith than most of us are willing to admit. The sad truth is that the extant arguments against dualism are not all that compelling.”[3] After explaining the severe weaknesses in standard objections to substance dualism, Bennett confesses, “But all told, we physicalists are perhaps not in as good a position as we like to think.”[4] In what follows, we intend to reveal even more weaknesses of anti-SD arguments.

10.1 THE THREE PROBLEMS OF CAUSAL INTERACTION

“The chief drawback of dualism,” says Jerry Fodor, “is its failure to account adequately for mental causation.”[5] In this section, we will state and respond to three objections raised against SD (and Property Dualists as well (PD).) In one way or another, each objection is an aspect of the problem of causal interaction.[6]

10.1.1 Disambiguating “How Can Mental Entities Causally Interact with Physical Entities?”

Sometimes an opponent of SD/PD (hereafter, SD) will simply press the question “How can mental entities causally interact with physical entities?” Dialectically, this move is supposed to place the SD in a bad place from which he/she will have a hard time escaping. Unfortunately, the question is ambiguous: it has two very different meanings.

First, the question may be a request for a mechanism between mental and physical entities in virtue of which they interact with each other. Call this the *Mechanistic Question*. For example, if one asks how turning on a car’s ignition switch causes the engine to start, one is requesting a description of a mechanism, a system that is causally between the ignition and the engine in virtue of which turning the former starts the latter.

If this is the correct interpretation, then for the SD, it is a complete non-starter. Why? Because the SD will hold that the causal connection between mental and physical entities is basic, there is nothing between them in virtue of which they causally interact with each other. Suppose the physicalist responds by saying that the discovery of such mechanisms is precisely what neuroscience gives us. In that case, the dualism will respond by pointing out that all neuroscience provides is increasingly precise correlations that leave mental/physical interaction basic even if the physical side of a correlation turns out to be a complicated system of neurons.

---

[2] Lycan, W. (2009). Giving dualism its due. *Australasian Journal of Philosophy*, *87*(4), 551.

[3] Bennett, K. (2021). Why I am not a dualist. In U. Kriegel (Ed.), *Oxford studies in philosophy of mind* (Vol. 1, p. 210). Oxford: Oxford University Press.

[4] Ibid., p. 212.

[5] Fodor, J. (1994). The mind–body problem. In R. Warner & T. Szubka (Eds.), *The mind–body problem* (p. 25). Oxford, UK: Blackwell.

[6] In what follows we will assume a *causal production view* of causation according to which a cause produces, brings about its effect when the cause’s powers are actualized. For more on this and other views of causation, see Swinburne, R. (2013). *Minds, brains, and free will* (pp. 125–140). Oxford: Oxford University Press.

Second, the question may be understood simply as an expression of skeptical exasperation, perhaps based on the differences between mental and physical entities. Call this the *Skeptical Question*. "Given these differences, how could such interaction take place? It strains credulity to believe in such a thing." As Keith Yandell notes, the skeptic seems to assume that causal interaction violates some necessary truth, such as (1) Like can only affect like; or (2) What is in space can only be affected by what is in space.[7]

Regarding (1), Yandell notes that it assumes the *Causal Likeness Principle*: If A causes Q in B, then A must itself have Q.[8] The problem with this principle is that there are numerous actual and possible counterexamples to it: retinal excitation produces color sensations, imbalance of ear fluids causes feelings of vertigo, the Big Bang produced massive mass emigration, space and time. In the history of science, particles were thought to interact with the ether, caloric, phlogiston, fields, and so on. Even if these theories turned out false, they could have been true; that is, they were metaphysically possible. Finally, we have far more grounds for believing that interaction takes place than we do for the *Causal Likeness Principle*. So, if anything must go, it is the latter, not the former.[9]

Regarding (2), propositions of the form

$A_{\text{(non-spatial)}}$ causally affects $B_{\text{(spatial)}}$, and
$B_{\text{(spatial)}}$ causally affects $A_{\text{(non-spatial)}}$,

do not entail Q and not-Q states of affairs.

$A_{\text{(non-spatial)}}$ may have Q as a causal disposition that it brings about in $B_{\text{(spatial)}}$ or conversely. These propositions could be true despite the spatiality of one of the causal relata and the non-spatiality of the other. Moreover, it is perfectly intelligible that God, even if there is no God, is omnipresent in a reductive sense. Strictly speaking, God is not in space, but each spatial location is that to which God has direct, unmediated causal and epistemic access and could part the Red Sea. What's the problem here? Finally, most SDs hold that the soul is spatially located (though most hold it is not spatially extended), so even if this objection is sound, it is irrelevant for most SDs.

10.1.2 Causal Interaction Violates the Conservation of Energy Principle

A basic principle of physics is that if the trajectory of any physical entity is changed, this amounts to an acceleration that requires the expenditure of energy. If one decides to raise one's arm to vote, then the sequence of physical events in the agent's brain and body goes a different direction (the arm goes up) than it would have (the arm stays down) if the agent had not acted on his/her decision. Thus, energy was expended and interjected into a physical system—the agent's brain and body. But this violates the Conservation of Energy Principle that energy can neither be created nor destroyed. Daniel Dennett refers to this as "the inescapable and fatal flaw with dualism."[10] According to John Searle, SD is incompatible with modern physics by assuming "another kind of energy, mental energy or spiritual energy, that is not fixed by physics."[11] Why is this a problem? Because the mental act intervenes into the brain/body and must create new energy since that intervention gives the system energy it did not have.

---

[7] Yandell, K. (1995). A defense of dualism. *Faith and Philosophy*, *12*, 551. Cf. 551–53. Yandell also mentions a third candidate: Ultimate connections cannot be brute. But this seems either incoherent—the meaning of "ultimate" seems to entail being brute strictly—or it is just plain false because physicalists allow for ultimate physical connections to be brute. Moreover, if such connections cannot be brute, this would strictly entail a vicious infinite regress.

[8] Ibid.

[9] See, e.g., Lowe, E. J. (2008). *Personal agency: The metaphysics of mind and action* (pp. 133–135). Oxford: Oxford University Press.

This objection is based on two assumptions: (1) The principle of conservation applies to all purely physical interactions; (2) All causal interactions between events must involve an exchange of energy. But as Robin Collins argues, "the first assumption is false for the case of general relativity ...; the second assumption is false for the case of quantum mechanics. Thus, based on current physics, the energy conservation objection has little, if any, merit."[12]

According to Collins, the principle of the conservation of energy is false in general relativity theory. A reference frame is basically a frame of reference relative to the location of an observer that is used for the observation and description of physical phenomena. Many physical phenomena, such as duration of time and spatial extension, are relative to a reference frame. But some physical entities are frame-independent quantities taken to be intrinsic characteristics of an object. In general relativity theory, a quantity known as the stress-energy tensor is such a frame-independent quantity of an object, but the object's energy cannot be so considered.

Thus, in relating the stress-energy tensor of an object to the conservation of energy, a problem arises: the tensor cannot be defined for gravitational fields or objects within them, so there is no law of conservation for interactions involving gravity. So, although gravitational fields causally influence objects in various ways, their influence cannot be understood concerning movement of energy through space. Within general relativity, energy is not conserved because the law of conservation is not well defined. Thus, in the presence of a gravitational field, the total non-gravitational energy within an enclosed region of space could increase/decrease without a corresponding net physically definable energy flowing across that region's boundary. So, there is a gain in energy without that energy existing somewhere beforehand. Not everyone accepts this line of argumentation about the conservation law and general relativity, and we are not qualified to weigh in on this debate. About all we can say is that there are respected scientists who agree with Collins's argument, and it seems to be a plausible one.[13]

Collins suggests that, like a gravitational field, the mind may well be a place where the conservation of energy simply does not apply. Thus, the mind could cause a real change in the brain's energy without energy being conserved.

---

[10] Dennett, D. (1991). *Consciousness explained* (p. 35). Boston: Little, Brown, and Company.

[11] Searle, J. (2004). *Mind: A brief introduction* (pp. 29–30). Oxford: Oxford University Press.

[12] Collins, R. (2011). A scientific case for the soul. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis* (p. 125). New York, NY: Continuum. We are indebted to Collins for the discussion that follows regarding these two claims.

[13] See Wald, R. (1984). *General relativity*. Chicago: University of Chicago Press; Hoefer, C. (2000). Energy conservation if GTR. *Studies in the History of Modern Physics*, *31*, 194; Robin Collins, endnote 15, 265–266.

Regarding (2), it is widely acknowledged that in a phenomenon known as *quantum entanglement*, causal interaction takes place without any exchange of energy. To grasp quantum entanglement, let us suppose we have a sub-atomic particle that has a physical attribute or quantum—e.g., spin—that can exist in only two ways, say up or down, and can only be described with respect to each other even if they are spatially distant from each other. For example, say our two particles are such that if one has an upward spin, the other must have a downward spin and conversely. Now, suppose our system includes these two particles. In that case, if one is measured to have an upward spin, the other will *simultaneously* be caused to have a downward spin even though there could not have been an exchange of energy due to the distance between the particles and the simultaneity of the causal effect. *Causation occurs without an exchange of energy, and there is no reason against the idea that mind/brain causal interactions do not involve an exchange of energy and, thus, do not violate the conservation of energy principle.*

Besides these two points, more can be said about the so-called conservation of energy problem.

1. Many have replied that the conservation objection assumes that SD is false. Others have argued that the objection cannot be clearly stated. If the conservation law is too strong, then it cannot be scientifically confirmed, but if it is too weak then it cannot exclude soul–body causation.[14]
2. The principle is not a metaphysical one, e.g., "necessarily, if something is colored, then it is extended" in which case it would be metaphysically necessary and true throughout all possible worlds. Rather, it is a mere empirical generalization like "all ravens are black" which could admit of exceptions. Moreover, the principle used to be stated properly: "In ordinary physical processes, energy is neither created nor destroyed." This intentionally left open the fact that mental causation was not an ordinary physical process and, thus, did not violate the conservation principle. If one responds that the principle is a scientifically accepted heuristic principle used to do research, then we simply reject the recommendation. And we also are not doing science in our defense of causal interaction.
3. Let us suppose that an exercise of active power involved in a free action does, in fact, involve the mental creation of energy. So what. The conservation principle does not even apply to mental causation. But, then why have we not detected the increase in energy in, say, the brain, that results from free action? We're not sure anyone has tried to look for such energy, but it is plausible that it may be so slight and insignificant, that it cannot be detected. Think of the massive amount of energy in the water held back by Hoover Dam. Now suppose someone flips a switch with his finger, and this opens up the dam releasing all that potential energy. The energy it took to flip the switch is real, but so small compared to the energy released by the dam opening that it would be undetectable. This may very well be what is going on with mental causation which triggers the release of comparatively large amounts of potential energy stored in the brain and body.

---

[14] See, Averill, E., & Keating, B. F. (1981). Does interactionism violate a law of classical physics? *Mind*, *90*(357), 102–107.

4. Perhaps SD does not actually violate the conservation of energy. C. D. Broad and Keith Campbell use the example of a pendulum: the pull of the string causally influences the way the weight moves and the distribution of energy (potential and kinetic) although the total energy of the weight remains the same.[15] According to Broad,

   > minds neither add energy to nor abstract it from the [body]. What they do, if they do anything, is to determine that at a given moment so much energy shall change from the chemical form to the form of bodily movement; and they determine this, so far as we can see, without altering the total amount of energy in the physical world.[16]

   The soul only redistributes energy, and so conservation is maintained.

5. Alin C. Cucu and J. Brian Pitts argue that, in order to be most consistent with contemporary physics (especially the physico-mathematical roots of energy conservation physics: the first Noether theorem), SDs should see energy conservation as (bi) conditional upon symmetries and accept energy non-conservation.[17] Additionally, Pitts argues that conservation of energy is local, not for the universe, but for everywhere separately. Global conservation holds only if the global summing-up of local conservation laws converges, and there are reasons to think it probably doesn't.[18]

For a more detailed response to problems with mental causation and the first and second conservation law, see the appendix written by our colleague Mihretu Guta.

10.1.3 Causal Interaction Falls Prey to the Problem of Causal Pairing

The dualist problem of Causal Pairing was first raised by Jaegwon Kim, then in more detail by Ernest Sosa and John Foster,[19] and has been revived and strengthened by Kim.[20] The argument is directed toward versions of Cartesian dualism that treat the soul as a non-spatially located entity causally connected to its body.

---

[15] Broad, C. D. (1925). *The mind and its place in nature* (pp. 107–108). London: Kegan Paul, Trench, Trubner, and Co.; and Campbell, K. *Body and mind* (2nd ed., pp. 53–53. Notre Dame: University of Notre Dame Press.

[16] Broad, p. 109.

[17] Cucu, A. C., & Pitts, J. B. (2019). How dualists should (not) respond to the objection from energy conservation. *Mind and Matter*, *17*(1), 95–121.

[18] See, Pitts, J. B. (2020). Conservation laws and the philosophy of mind: Opening the black box, finding a mirror. *Philosophia*, *48*(2), 673–707; and Pitts, J. B. (2022). General relativity, mental causation, and energy conservation. *Erkenntnis*, *87*(4), 1931–1973.

[19] See, Kim, J. (1973). Causation, nomic subsumption, and the concept of event. *Journal of Philosophy*, *70*, 217–236; Sosa, E. (1984). Mind-body interaction and supervenient causation. *Midwest Studies in Philosophy*, *9*, 271–281; and Foster, J. (1991). *The immaterial self: A defense of the Cartesian dualist conception of the mind* (pp. 163–172). London: Routledge.

[20] See Kim, J. (2005). *Physicalism, or something near enough* (pp. 78–90). Princeton, NJ: Princeton University Press.

Suppose we have two physical objects, A and B, e.g., rifles, that are physically indistinguishable in all respects except for their spatial location. Further, there are two targets similarly indistinguishable. When A and B fire a shot simultaneously, why do A and B hit targets C and D, respectively? The answer seems obvious: Focusing on A, it is pointed at C, not D, and there is a spatially continuous causal path from A to C. In general, says Kim, it is this spatial location, orientation of causes and effects, and a properly directed causal pathway between them that explains why the cause brought about the effect.

But now a problem arises for the version of dualism under consideration. Suppose two people are in mentally indistinguishable conscious states (e.g., willing to raise one's left arm), and two bodies/brains are physically indistinguishable. Why does person A's mental state cause an effect in one body (e.g., the left arm moving upward) rather than the other one? The answer cannot be that A's mental state causes *his* arm to move, not the other person's, because this answer is circular. Why? A causes his arm to move, but what makes that specific arm A's is that A's soul is causally connected to it. Put simply; A causes the arm to move with which A is causally connected. This is circular. Nor can the answer be that A and his mental state are spatially oriented toward one arm and not the other as in the rifle case because A's soul and mental states are not spatially located. It seems, then, that mental/physical causal interaction is incoherent.

In response to Kim's claim that *causality requires pairing relations, and there are no such relations for immaterial minds*, Greg Ganssle provides three compelling responses:

> First one could deny the first part of the claim and argue that causality does not need a pairing relation. If causation in general does not need a pairing relation, then, perhaps there is no need for a pairing relation in cases of mental causation. Second, one could agree that there is a need for a pairing relation and that the only viable kind of pairing relation is spatial. One could then hold that souls are themselves spatial. Thus a spatial relation *can* provide the right kind of link when a mind acts on a particular state of affairs. Third, one can argue that, even if minds are not spatial and a pairing relation is needed, there is another relation that will work.[21]

Before we look at examples of these three strategies, Kim's causal-pairing argument, especially its claim that any solution to it requires a spatial location and proper orientation for a cause and its effect, bumps up against an insufferable rebutting defeater: It entails that the singularity of (the still widely accepted) standard model of the Big Bang could not have a cause since the model implies that the entire space–time continuum began at the singularity. Even if the standard model turns out to be false, surely it is coherent and metaphysically possibly true. To avoid this defeater, it seems like one must reject the principle of *ex nihilo nihil fit*. In our view, it is hard to take such a rejection seriously. As Kim has admitted, physicalism cannot be had on the cheap. Unfortunately, this way of avoiding the Big Bang defeater is purchased at a far too expensive price.

Returning to the strategies mentioned earlier, as an example of the first strategy, Andrew Bailey, Joshua Rasmussen, and Luke Van Horn provide a defeater for the claim that there must be some relation to link cause and effects.[22] Their argument centers on the idea that the sort of agency held by many dualists rules against it:

---

[21] Ganssle, G. (2022). Divine causation and the pairing problem. In G. E. Ganssle (Ed.), *Divine causation* (p. 6). London: Routledge.

> Now consider a world in which two persons Tim and Tom, are exactly similar in all respects (excluding, of course, their haecceitous properties). Suppose that Tim and Tom each have the same two options available to them – to cause *A* or to refrain from causing *A*. If Tom happens to cause *A* while Tim refrains, then we have a situation in which [the general requirement for a pairing relation between cause and effect] fails.[23]

They conclude that Kim's argument works only against versions of dualism that do not include an agent-causation view. Any position that involves this kind of agency will be immune from Kim's argument. Since we embrace SLA, their defeater applies to our position.

Greg Ganssle points out another version of this first strategy:

> Another consideration is that singular causal theories, such as the one defended by Michael Tooley, do not require a general causal relation to pair causes and effects: "Yet if singular causal relations are possible then it seems that it should be possible for a cause to stand in a singular causal relation to an effect while there is an indistinguishable candidate cause in the neighborhood, given that singular causal relations are not fixed/determined by the properties that a thing has."[24] Theories that analyze causation along these lines do not understand causation in terms of any general principle at all. A particular object may stand in a causal relation to another without that relation being subject to analysis in terms of distinguishable characteristics the object may have.[25]

An example of the second strategy involves the observation that virtually all contemporary substance dualists believe the soul is spatially located even though it is not extended. Some provide arguments for the soul being located in one specific place in the body.[26] Others believe that the soul is to the body as God is to space, namely, fully present at each location in the body. This view is called holenmerism.[27] Either way, the spatial location requirement is satisfied.

---

[22] Bailey, A. M., Rasmussen, J., & Van Horn, L. (2011). No pairing problem. *Philosophical Studies*, *154*, 349–360. They reproduce Kim's argument as follows:

1. Necessarily, for all *x* and all *y*, if *x* causes *y*, then there is a relation or relations, *R*s, such that their holding makes it the case that *x* cause *y*.
2. Necessarily, a spatial relation is among the *R*s.
3. Necessarily, a soul stands in no spatial relations.
4. Therefore, it is impossible that a soul causes anything.

We are indebted to Greg Ganssle for pointing this out to us.

[23] Bailey, et al., No pairing problem, p. 354.

[24] Ibid., p. 355.

[25] Ganssle, p. 7.

[26] See Meixner, U. (2010). Materialism does not save the phenomena—and the alternative which does. In R. C. Koons & G. Bealer (Eds.), *The waning of materialism* (pp. 417–437). Oxford: Oxford University Press.

[27] See Moreland, J. P. (2018). In defense of a Thomistic-like dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 102–122). Oxford: Wiley-Blackwell; Rozemond, M. (2003). Descartes, mind-body union, and holenmerism. *Philosophical Topics*, *31*(1/2), 343–367; and Skirry, J. (2001). A hylomorphic interpretation of Descartes's theory of mind-body union. *Proceedings of the American Catholic Philosophical Association*, (75), 267–283.

As an example of the third approach, assuming the soul is not located in space, there may be a metaphysical analog to spatial location that solves the problem. Suppose numbers are real abstract objects: 1, 2, ... There is a real sense in which each number is "located" in a specific "place" in the number sequence. Similarly, there may be a metaphysical grid in which souls are "located." If this is true, then soul A causes movement in a particular body because A is properly metaphysically located concerning a particular body, which is what makes that body A's. Accordingly, A then causally interacts with and only with *his/her* body.

According to another version of this third approach that is consistent with the soul being or not being located in space, some dualists provide a deeper grounding than causation for what makes a given body belong to a specific soul. As Foster pointed out, the dualist can specify some

> relation which holds between any mental event in some subject's mind and any physical event in the *same* subject's body—in other words, the relation signified by the expression "x belongs to a subject in whose body y occurs." Let us, for convenience, abbreviate this expression to "x is subject-linked to y".[28]

For example, Thomistic dualists claim that the soul (or its essence or a substantial form) informs a body, and that is what makes that body belong to a particular soul. Then the Thomist can say that a soul causally interacts directly with, and only with, its own body. Emergent substance dualists hold that the soul interacts directly with and only with the brain (and the body having that brain) from which the soul emerges.

Finally, Foster has offered a response to the problem, which, if successful, seems to imply that dualist and physicalist problems with causal pairing cancel each other out.[29] The details of Foster's thought experiment are interesting but irrelevant to our point. That point is this: If successful—and it certainly seems to be—Foster shows that physicalists, and dualists, face the same causal pairing problem. Thus, physicalists must provide a solution to this problem every bit as much as dualists should, and we have attempted to do just that. However, given the parity of this problem, it can hardly be proffered as a difficulty unique to dualism.[30]

More could be said in response to the Problem of Causal Pairing, but enough has been offered to show that the problem is hardly a defeater for mental/physical causal interaction.[31]

---

[28] Foster, *The immaterial self*, p. 165.

[29] Ibid., pp. 170–171.

[30] See also Bennett, K. (2007). Mental causation. *Philosophy Compass*, *2*(2), 320–322; and Saad, B. (2018). A causal argument for dualism. *Philosophical Studies*, *175*(10), 2475–2506.

[31] Addressing the problem of the causal closure of the physical is beyond the scope of this short chapter. For a powerful statement of the argument, see Kim, *Physicalism, or something near enough*, pp. 13–22, 32–69. For dualist responses to the argument, see Moreland, J. P. (2016). Why top-down causation does not supply adequate support for mental causation. In T. M. Crisp, S. L. Porter, & G. A. Ten Elshoff (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology* (pp. 51–73). Grand Rapids, MI: Eerdmans; Lowe, E. J. (2010). *Personal agency* (pp. 41–78). Oxford: Oxford University Press; Hasker, W. (1999). *The emergent self* (pp. 58–80). Ithaca, NY: Cornell University Press; Goetz, S., & Taliaferro, C. (2011). *A brief history of the soul* (pp. 159–176). Malden, MA: Wiley-Blackwell.

10.2 NINE NEUROSCIENTIFIC OBJECTIONS TO SD

In this section, we will respond to nine objections against SD that involve neuroscientific practices and discoveries in one way or another. A line from Roger Scruton captures the general problem: "Recent advances in genetics, neuroscience, and evolutionary psychology have all but killed off that idea [substance dualism]."[32] "The reason for this radical change of opinion [the rejection of the soul]," says Francis Crick, "is the spectacular advance of modern science."[33] More pointedly, William Seager argues that

> the idea that the property of being conscious applies only to metaphysically simple entities appears to endorse a kind of dualism closely akin to that of Descartes and as such is not a very attractive option. It seems to be entirely at odds with the huge amounts of evidence we already possess that consciousness is intimately associated with the brain's large-scale dynamical and structural features of the brain. It really does "look" as if consciousness is the large-scale product and property of the brain or nervous system.[34]

The objection is that substance dualism is somehow at odds with neuroscience. However, it isn't clear what this amounts to. Is this a soul-of-the-gaps objection, or is this a defeater based on neuroscientific discoveries? Perhaps it is both.

10.2.1 SD as a Soul-of-the-Gaps Argument

A soul-of-the-gaps objection fails to understand why most dualists hold their view. Plato, Augustine, Descartes, Butler, and Reid held dualism in virtue of being aware of themselves from the first-person perspective as not reducible to or identical to their body.[35] Others, like Aristotle and Aquinas, arrived at different kinds of dualism by analyzing positive arguments for the soul. After a detailed look at the literature, Stewart Goetz and Charles Taliaferro conclude, There is not the least bit of evidence for the idea that they arrived at their belief in the soul's existence after failing to explain various experiences in terms of what goes on in the physical world.[36]

The same is true of contemporary dualists. We've yet to find a single contemporary dualist who posits the soul to fill an explanatory scientific gap. It is the awareness of self and metaphysical arguments that motivate dualism. This accords well with the prevailing view among cognitive scientists that dualism is a widespread, pre-theoretical belief

---

[32] Scruton, R. (2017, March 6). If we are not just animals, what are we? *New York Times*. https://mobile.nytimes.com/2017/03/06/opinion/if-we-are-not-just-animals-what-are-we.html?referer=http%3A%2F%2Fwww.thepublicdiscourse.com%2F2017%2F05%2F19297%2F.

[33] Crick, F. (1994). *The astonishing hypothesis: The scientific search for the soul* (p. 4). New York, NY: Charles Scribner's Sons.

[34] Seager, W. (2010). Panpsychism, aggregation, and combinatorial infusion. *Mind and Matter*, *8*(2), 174.

[35] See our arguments in Chapter 5.

[36] Goetz, S., & Taliaferro, C. (2011). *A brief history of the soul* (p. 155). Oxford: Wiley-Blackwell.

(including non-religious),[37] shared across cultures, and developed in infancy.[38] We are, as Paul Bloom says, “natural Cartesians.”[39]

10.2.2 Neuroscience and the Explanatory Impotence of SD

In an important critique of SD filled with a number of objections, Ian Ravenscroft asserts the following:

> In this subsection, I will advance what I take to be *the most serious* objection to SD—the observation that physicalist theories of the mind have vastly more explanatory power than SD.[40]

Along similar lines, Tim Bayne complains that with respect to a wide range of problems regarding consciousness, while physicalism has a solid track record of explaining such problems and offers the prospect of accounting for further difficulties, “[s]ubstance dualists offer no answer to these questions.”[41]

Finally, Nancey Murphy opines that “science has provided a massive amount of evidence suggesting that we need not postulate the existence of an entity such as a soul or mind in order to explain life and consciousness.”[42] This evidence consists of the fact that “biology, neuroscience, and cognitive science have provided accounts of the dependence on physical processes of *specific* faculties once attributed to the soul.”[43] Elsewhere, she claims:

> My argument in brief is this: all of the human capacities once attributed to the mind or soul are now being fruitfully studied as brain processes—or, more accurately, I should say, processes involving the brain, the rest of the nervous system and other bodily systems, all interacting with the socio-cultural world.[44]

---

[37] See, e.g., Bering, J., Blasi, C. H., & Bjorklund, D. (2005). The development of “afterlife” beliefs in religiously and secularly schooled children. *British Journal of Developmental Psychology*, *23*, 587–607; Bering, J., & Bjorklund, D. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, *40*(2), 217–233.

[38] See, e.g., Bloom, P. (2004). *Descartes’ baby: How the science of child development explains what makes us human*. Cambridge, MA: Basic Books; Kuhlmeier, V. A., Bloom, P., & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition*, *94*(1), 95–103; Wellman, H., & Hickling, A. K. (1994). The minds ‘I’: Children’s conception of the mind as an active agent. *Child Development*, *65*, 1564–1580; Wellman, H. (1990). *The child’s theory of mind*. Cambridge, MA: MIT Press; and Wellman, H., & Estes, D. (1986). Early understanding of mental entities: A re-examination of childhood realism. *Child Development*, *57*, 910–923.

[39] Bloom, *Descartes’ Baby*, xii. For a philosophical analysis of this literature in relation to dualism, see, Rickabaugh, B. (2019). Alister McGrath’s anti-mind-body dualism: Neuroscientific and philosophical quandaries for Christian physicalism. *Trinity Journal*, *40*, 215–240.

[40] Ravenscroft, I. (2018). Why reject substance dualism. In J. J. Loose, A. J. L. Menuge & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 267–282). Oxford: Wiley-Blackwell, especially 276–278. The quote is from 276. Italics ours.

[41] Bayne, T. Problems with unity of consciousness arguments, in Ibid., 208–225. The quote is from 223.

[42] Murphy, N. (1998). Human nature: Historical, scientific and religious issues. In W. S. Brown, N. Murphy, & H. Newton Malony (Eds.), *Whatever happened to the soul?* (p. 18). Minneapolis, MN: Fortress Press.

[43] Ibid, p. 17; cf. 13, 27.

[44] Murphy, N. (2006). *Bodies and souls, or spirited bodies?* (p. 56). Cambridge: Cambridge University Press.

Murphy acknowledges that “dualism cannot be *proven* false—a dualist can always appeal to correlations or functional relations between soul and brain/body—but advances in science make it a view with little justification.”[45]

What should one say about this line of reasoning? We offer five responses. First, many substance dualists do not believe in a substantial ego primarily because it is a theoretical postulate with superior explanatory power. Instead, they take the ego to be something of which people are directly aware. As Walter Hopp succinctly puts it regarding the phenomenology of consciousness: “For many … approaches to the mind, conscious experiences are theoretical postulates postulated to explain certain observable data. For the phenomenologist, they are the data.”[46] For many SDs, Hopp’s point applies with equal force to the soul. We take William Lycan’s more general statement to be canonical: “Cartesian minds are not explanatory posits at all, much less posited to explain physical facts. They are known from the inside and there is nothing queer or obscure about that.”[47]

Thus, belief in a substantial, simple soul is properly basic and grounded in self-awareness and first-person reports. The point is not that dualists are right about this, though we have supported this approach in Chapter 5. Rather, given this dualist approach, the point is that advances in our knowledge of mental/physical dependencies are simply beside the point. Dualist entities—conscious states and the soul—are not embraced primarily due to their theoretical or explanatory superiority. They are entities of which we have direct acquaintance. And the further debate about which approach (the direct awareness or the theoretical postulate approach) is the fundamental one for defending SD is not something for which advances in scientific knowledge are relevant. For more on this, see Chapter 4.

Second, in those cases where SD (or a substantial form) *is* postulated as the best explanation for a range of purported facts, typically, those facts are distinctively philosophical and seldom the scientific ones Murphy mentions. This entire book illustrates this fact. Arguments from the unity of consciousness, the possibility of disembodied survival or body switches, the best view of an agent to support libertarian agent causation, the metaphysical implications from the use of the indexical “I,” and the special sort of diachronic and synchronic unity of human persons are typical of arguments offered by substance dualists. The fact that SD does not explain the sort of facts treated by empirical neuroscientific theory is beside the point. SD is a philosophical position not primarily proffered to explain those facts.

Murphy’s position about the scientific nature of SD is not benign. It implies a distorted, highly misleading picture of the debate between SDs and their physicalist counterparts. To cite one example, when Paul Churchland sets up the criteria for adjudicating the debate between dualist and physical rivals, he begins by noting that these are different explanatory theories and, thus, we need epistemic criteria for assessing them. For this task, Churchland offers these criteria: greatest explanatory power, predictive power, coherence, and simplicity.[48] Predictive power? What is that doing in the list? Greater explanatory power? Churchland employs this criterion in a way that is significantly directed toward explaining the scientific facts about reduction, brain processes, and the like. Similarly, coherence is spelled out in terms of cohering with our generally accepted scientific facts, and simplicity is used as a razor to shave off any dualist entities, given that different forms of reduction make them otiose. When we read Churchland, it became clear that he stacked the deck from the beginning to favor physicalism by setting the terms of the debate in virtue of stating and employing criteria for theory selection, criteria that may be appropriately suited for assessing scientific theories but are misleading and question-begging in a philosophical dialog.

---

[45] Murphy (2006), p. 112.

[46] Hopp, W. (2011). *Perception and knowledge* (p. 4). Cambridge: Cambridge University Press.

[47] Lycan, W. (2018). Redressing substance dualism. In J. Loose, A. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (p. 27). Oxford: Wiley-Blackwell.

[48] Churchland, P. M. (2013). *Matter and consciousness* (3rd rev. ed., p. 11). Cambridge, MA: MIT Press, cf. 12–35.

Third, some dualists supplement their philosophical arguments for SD by secondarily bolstering their view in light of neuroscientific data.[49] Even non-dualists recognize that neuroscientific methods are based on dualism's conceptual framework, as they take first-person reports as epistemically privileged over and, in fact, required for selecting the relevant third-person observations.[50] More on this last point below.

Fourth, others argue that neuroscience fails to support physicalism over dualism.[51] If one is to defend the conflict between neuroscience and dualism, one must offer plausible replies to these last two lines of argument. Such replies are, as of yet, absent.

Finally, the discovery of "the dependence on physical processes of *specific* faculties once attributed to the soul" does not provide sufficient grounds for attributing those faculties to the brain rather than to the soul. There is an important distinction between describing the nature, proper categorization, and possessor of a capacity vs. explaining what conditions are necessary for its actualization. To see this, it is important to get clear on the use of "faculty" as the term has been historically used in discussions of substances in general and the soul in particular.[52] Roughly, a faculty of some particular substance is a natural grouping of resembling capacities or potentialities possessed by that thing. For example, the various capacities to hear sounds would constitute a person's auditory faculty.

Moreover, a capacity gets its identity and proper metaphysical categorization from the type of property it actualizes.[53] The nature of a capacity-to-exemplify-F is properly characterized by F itself. Thus, the capacity to reflect light is properly considered a physical, optical capacity. For property dualists, the capacities for various mental states are mental and not physical capacities. Thus, the faculties constituted by those capacities are mental and not physical faculties.

---

[49] See Manzotti, R., & Moderato, P. Neuroscience: Dualism in disguise. In Robinson and Lavazza, *Contemporary dualism* (pp. 81–97); Antonietti, A. Must psychologists be dualists? In Antonietti, Corradini, & Lowe, *Psycho-physical dualism today* (pp. 37–67).

[50] See, e.g., Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell; Rockwell, W. T. (2005). *Neither ghost nor brain*. Cambridge, MA: MIT Press; Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press, and Uttal, W. R. (2004). *Dualism: The original sin of cognitivism*. Mahwah, NJ: Lawrence Erlbaum Associates.

[51] See, e.g., LaRock, E. (2016). Neuroscience and the hard problem of consciousness. In T. Crisp, S. Porter, & G. Ten Elshof (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology* (pp. 151–180). Grand Rapids, MI: Eerdmans; LaRock, E., & Collins, R. Saving our souls from materialism. In *Neuroscience and the soul* (pp. 137–146); LaRock, E. (2008). Is consciousness really a brain process? *International Philosophical Quarterly*, *48*(2), 201–222; Schwartz, J. M., & Begley, S. (2002). *The mind and the brain: Neuroplasticity and the power of mental force*. New York, NY: HarperCollins; Guta, M. P. (2011). Neuroscience or neuroscientism? *Perspectives on Science and Christian Faith*, *63*(1), 69–70; and Horgan, T. (1993). Nonreductive materialism and the explanatory autonomy of psychology. In S. Wagner & R. Warner (Eds.), *Naturalism: A critical appraisal* (pp. 313–314). Notre Dame, IN: University of Notre Dame Press.

[52] See Perler, D. (2015). *The faculties: A history*. Oxford: Oxford University Press.

[53] For multi-tracked dispositions, an actualized manifestation characterizes an essential constituent of the disposition's essence but is not exhaustive of it. We are using powers and dispositions interchangeably.

Now, arguably, a particular is the kind of thing it is by virtue of the actual and potential properties/faculties essential and intrinsic to it. Thus, a description of the faculties of a thing provides accurate information about the kind of particular that has those faculties. Moreover, a description of a particular's capacities/faculties is a more accurate source of information about its nature than is an analysis of the causal/functional conditions relevant for the particular to act in various ways. The latter can either be clues to the intrinsic nature of that particular or else information about some other entity that the particular relates to in exhibiting a particular causal action. Remember, there is a difference between attempts to describe, categorize, and identify a capacity's nature and possessor as opposed to proffering an explanation of the functional/causal conditions that must be present for that capacity to be actualized.

For example, if Smith needs to use a magnet to pick up certain unreachable iron filings, information about the precise nature of the magnet and its role in Smith's action does not tell us much about the nature of Smith (except that he is dependent in his functional abilities on other things, e.g., the magnet). We surely would not conclude that a magnet's actual and potential properties are clues to Smith's inner nature. Similarly, functional dependence on/causal relations to the brain are of much less value in telling us what kind of thing a human person is than a careful description of the kind-defining mental capacities, i.e., faculties, human persons as such possess.

10.2.3 Empirically Equivalent Theories and the Findings of Neuroscience

The neuro-scientific empirical discovery of mental/physical correlations (causal or dependency relations both ways) provides no evidence whatsoever for physicalism. Indeed, as we will see, the central issues regarding the mind—what is a thought, feeling, or belief; what is that to which one's self is identical—are basically common sense and philosophical issues for which scientific discoveries are largely irrelevant. Science helps answer questions about what factors in the brain and body generally hinder or cause mental states to obtain. Still, science is largely silent about the nature of mental properties/states.

To see the basic irrelevance of empirical data, consider the following. We have discovered that if a certain type of neuron—*mirror neurons*—is damaged, then one cannot feel empathy for another. How are we to understand this? To answer this question, we need to get the notion of empirically equivalent theories before our minds. If two or more theories are empirically equivalent, then they are consistent with all and only the same set of empirical observations. Thus, an appeal to empirical data cannot be made in favor of one of such theories over the others.

Three empirically equivalent solutions to the discovery of the function of mirror neurons come to mind: (1) strict physicalism (a feeling of empathy is identical to something physical, e.g., the firings of mirror neurons); (2) mere property dualism (a feeling of empathy is an irreducible state of consciousness *in the brain* whose obtaining depends on the firing of mirror neurons); (3) substance dualism (a feeling of empathy is an irreducible state of consciousness *in the soul* whose obtaining depends (while embodied) on the firing of mirror neurons).

Of these three, no empirical datum can pick which is correct, nor does an appeal to epistemic simplicity help. This is why three Nobel Prize winners working in neuroscience and related fields—John Eccles (substance dualist), Roger Sperry (mere property dualist), and Francis Crick (strict physicalist) could hold different ontologies regarding consciousness and the self even though they all knew the same neuroscientific data. Epistemic simplicity is a tie-breaker, and the substance dualist will insist that the arguments and evidence for substance dualism are better than those for the other two options mentioned above.

For at least two reasons, it does not follow from these points that empirical data are never relevant to building a case for one metaphysical view vs. another, e.g., property or substance dualism vs. strict physicalism (everything is physical). First, we can distinguish between *strong* and *weak empirical equivalence*. Applied to issues in philosophy of mind, the former implies that it is never the case, or at least, it is not the case for a particular issue, that empirical data themselves provide any grounds at all in favor of a physicalist vs. a dualist position (e.g., as in mirror neurons). In contrast, the latter implies that there can be some issues where a physicalist and dualist view are logically consistent with the empirical data. However, those data can still contribute to a cumulative case for one view vs. the other. For example, in previous generations, an extreme version of substance dualism implied no dependency or interaction with the brain at all! Empirical studies have shown extreme substance dualism to be false. Only a case-by-case basis can decide the relevance of empirical data to a mind/body debate.

Second, though property or substance dualism are fundamentally philosophical positions, in a secondary sense, they can generate a research program and make predictions. For example, by arguing that information is *sui generis*, irreducible, and grounded in minds, with auxiliary assumptions, one can predict that the role information plays in the origin and development of kinds of organisms is central and uneliminable. Given certain substance dualist views of the soul, it follows that libertarian free will is true, that consciousness is a holistic unity and not capable of being analyzed atomistically, and that there should be a binding problem not suggested by physicalist views of the "self," that there should be strong evidence for Near Death experiences, and so forth. Moreover, some versions of substance dualism entail a holistic organicism, an irreducible complexity, regarding organism's bodies.

In an important book endorsed by Christof Koch—*Measuring the Immeasurable Mind*—Matthew Owen claims that the book

> "argues that the possibility of measuring consciousness, with its corresponding medical benefits, is not ruled out by a dualist human ontology that disavows the doctrine of physicalism … [My] model gives a metaphysical explanation of NCC [the neural correlates of consciousness] informed by Aristotle's understanding of causation and a substance dualist human ontology inspired by the thought of Thomas Aquinas … A practical implication is that consciousness researchers need not subscribe to physicalism to be rationally justified in their search for NCC and a practical means for empirically discerning and measuring consciousness. Physicalists *and* dualists can co-labor on one of the most practically pressing issues in consciousness research."[54]

---

[54] Owen, M. (2021). *Measuring the immeasurable mind: Where the contemporary neuroscience meets the Aristotelian tradition* (pp. ix–x). Lanham, MD: Lexington Books.

10.2.4 Neuroscience and the Nature of the Central Issues in Philosophy of Mind

Upon careful reflection upon the central issues in philosophy of mind, including those in this book, it becomes evident that the methods and findings of neuroscience are simply irrelevant in addressing and solving these questions. These questions can be grouped into five areas of reflection.

> *Ontological Questions*: To what is a mental or physical property identical? To what is a mental or physical event identical? To what is the owner of mental properties/events identical? What is a human person? How are mental properties related to mental events (e.g., Do the latter exemplify or realize the former?)? Are there (Aristotelian or Leibnizian) essences and, if so, what is the essence of a mental event or of a human person? Am I essentially a human, a person, or both? What is free will?
>
> *Epistemological Questions*: How do we come to have knowledge or justified beliefs about other minds and about our own minds? Is there a proper epistemic order to first-person knowledge of one's own mind and third-person knowledge of other minds? How reliable is first-person introspection and what is its nature (e.g., a non-doxastic seeming or a disposition to believe)? If reliable, should first-person introspection be limited to providing knowledge about mental states or should it be extended to include knowledge about one's own ego?
>
> *Semantic Questions*: What is a meaning? What is a linguistic entity, and how is it related to a meaning? Is thought reducible to or a necessary condition for language use? How do the terms in our commonsense psychological vocabulary get their meaning? What is (or are) the Private Language argument(s), and are there adequate defeaters for it (them)?
>
> *Methodological Questions*: How should one proceed in analyzing and resolving the first-order issues that constitute the philosophy of mind? What is the proper order between philosophy and science? Should we adopt some form of philosophical naturalism, set aside so-called first philosophy, and engage topics in philosophy of mind within a framework of our empirically best-attested theories relevant to those topics? What is the role of thought experiments in the philosophy of mind, and how does the "first-person point of view" factor into generating the materials for formulating those thought experiments? Is conceivability a reliable guide to modality? What is the difference between epistemological and ontological simplicity, and what are their roles in the philosophy of mind?
>
> *Broad Worldview Issues*: What are the worldview implications of the existence of irreducible conscious states? Of an irreducible soul? Is theism or some version of panpsychism a better explanation for irreducible mental entities? What is the difference between staunch and faint-hearted naturalism, and which one should a naturalist embrace?

If one reflects carefully on these questions, it becomes clear that they are distinctively philosophical (and theological) questions for which neuroscience is significantly irrelevant. The simple fact is that scientific practitioners in neuroscience and related fields have virtually no training or exposure to philosophy of mind. UCLA neuroscience professor Daniel Siegel studied nearly eighty-thousand mental health practitioners from numerous countries on four continents, including psychologists and psychiatrists. He discovered that only 2 to 5% of the people in those fields had heard even one lecture on philosophical issues about the mind.[55] Is it any wonder why, when scientists in the relevant fields make claims about issues constituting philosophy of mind, those assertions are frequently misguided, uninformed, and taken with a grain of salt?

Nevertheless, some philosophers reject our claim that these questions are distinctively philosophical by arguing that they are examples of so-called First Philosophy, and for certain reasons, First Philosophy should be rejected.

As we have surveyed the literature on the rejection of First Philosophy, most "arguments" against it merely start with a question-begging assumption of weak or strong scientism. According to David Papineau, for example, if we are going to study the nature of the human mind, we must start with assumptions. (Note: As we pointed out in Chapter 5 and earlier in this chapter, we don't think we (epistemically) start with assumptions in tackling the core issues listed above; rather, we begin with our direct, introspective awareness of our conscious states.) "And the obvious strategy, naturalists will argue, is to begin with our empirically best-attested theories of the mind and its relation to reality, and use these as a framework within which to raise and resolve philosophical difficulties."[56] As we read Papineau, he offers very little by way of argument for this scientistic starting point. He just assumes it.

Moreover, when asked about the proper way to resolve disputes about the nature of knowledge or the best way to approach the meta-issue regarding what first-order methods should be employed—e.g., those listed in (4) above—Papineau asserts that he will

> conduct my argument within the framework of various empirical assumptions about the nature and needs of human beings. My defense of this strategy is that the onus surely lies with those who want to exclude relevant and well-confirmed empirical claims from philosophical debate to provide some rationale for doing so.[57]

The crucial claim is that one should not exclude relevant empirical claims from philosophical debate, but whether empirical claims are relevant is precisely what is at issue. More on this below.

In our view, the most widely stated argument against First Philosophy is that those who accept it do so because they hold to a faulty view of knowledge known as Cartesian foundationalism, which entails that knowledge requires certainty. As Patricia Churchland proclaims,

> It is in this sense that there is no first philosophy. There is no corpus of philosophical doctrine concerning science and epistemology such that we can be sure it is the Truth to which all science must conform. There is ... no Archimedean point outside all science from which we can pronounce upon the acceptability of scientific theories.[58]

---

[55] Siegel, D. (2010). *Mindsight: The new science of personal transformation* (pp. 50–51). New York, NY: Random House.

[56] Papineau, D. (1993). *Philosophical naturalism* (p. 3). Oxford: Blackwell.

[57] Ibid., p. 4.

[58] Churchland, P. (1986). *Neurophilosophy: Toward a unified science of the mind/brain* (p. 265). Cambridge, MA: MIT Press.

Papineau says pretty much the same thing. Speaking of the idea that philosophy needs firmer foundations than science and, indeed, without these firmer foundations, philosophy cannot justify the assumptions of science; he says,

> [T]his argument depends on the assumption that knowledge needs to be certain in the sense that it should derive from methods that necessarily deliver truths … If you hold that knowledge requires certainty, then you will hold that philosophy needs to come before science.[59]

If you look at the statements by Papineau and Churchland above as to why some philosophers accept First Philosophy, they both think that it is because those philosophers are Cartesian foundationalists. The idea is that if the assumptions of science are not completely certain and totally immune from error, then science is without adequate justification, and a skeptic about the claims of science wins.

Unfortunately for Churchland, Papineau, and others who use this argument, it is entirely wrongheaded and over fifty years outdated. Nothing could be further from the truth. Today, most epistemologists are foundationalists, but we do not know of one who is a Cartesian foundationalist. Cartesian foundationalism was given up over fifty years ago. Today, foundationalists believe that certain experiences can serve as solid internalist foundations for our properly basic beliefs without those experiences being immune from error or possibly being produced in a deviant way. So while an experience of *seeming* to hear something does justify the properly basic belief that one does, in fact, hear it, the experience is defeasible.

Papineau and Churchland are seriously mistaken as to why many scholars still accept First Philosophy. Why do these scholars do so? The answer turns out to be pretty obvious. When we examine cases that are purportedly *autonomous* from science (outside the competency of science), such as the proper system of ethics, the nature of modal logic, and so forth, it becomes obvious that the basic issues in these areas of reflection have nothing to do with physics, chemistry, and so forth. The basic issues are not stated in scientific language nor could debates about them be resolved by a scientific experiment.

When we examine cases that are purportedly *authoritative* with respect to science (the philosophical considerations carry more weight than the scientific ones), often, we can easily decide whether the case is or is not a genuine example of the authority of philosophy. All one needs to do is to state the different philosophical views of the issue and the key arguments involved, and do the same for the various scientific views and arguments.

Next, see if one cannot fully resolve the scientific issues until the philosophical ones have been weighed. If that is true, then we have an example of the authority of philosophy. If the scientific issues need to be resolved before we can consider the philosophical ones, then we do not have an example of the authority of science. We judge this on a case-by-case basis. We think that the five categories of topics listed above are distinctively philosophical, and we ask you, the reader, to return to them and see if you agree.

10.2.5 Neuroscientific Methodology Relies on First-Person Reports

In establishing mental/physical correlations (causal or dependency relations), neuroscience must assume the commonsense dualist notion of private access. In studying brain events to establish these correlations, the relevant mental state is made aware to the subject by first-person introspection constituted by direct, private access by that subject. These neuroscientific correlations ultimately depend on first-person reports regarding what is going on in their consciousness, just as the dualist predicts. But it is extremely odd and inexplicable that this would be necessary if mental states and the self were physical. If this were the case, all the facts involved in such studies would be publicly accessible, but they are not.

---

[59] Papineau, p. 4.

Moreover, as Richard Swinburne has noted, given the fact that the neuroscientist must rely on first-person testimony to know that a particular conscious event occurred, "if apparent testimony is to constitute evidence that conscious events occurred, the scientist must ... assume that the *subject* is caused to say what they do by a *belief* that the conscious events occurred and an *intention*, to *tell the truth* about their *belief*."[60] [italics ours] The terms in italics refer to mental entities. If we do not assume this and, instead, take them as physical, then the "subject's report" is nothing but sounds caused by brain events that cause them. In this case, physical determinism obtains. One would no more trust a subject's "report" than the sounds that come out of a parrot's mouth.

10.2.6 Scientific Explanation, Necessitation, and the Failure of the Neuroscientific Correlations

In standard scientific explanation, a state (also called an event) or a change of state is explained. For example, we may try to explain why the pressure of a gas is in a certain state or why it changed into that state. Or we might try to determine the reason a tornado was formed. In all these cases, one event causes another event in accordance with some law of nature. So, we have *event–event causation* (event A causes event B) or *state–state causation* (state A causes state B).

Associated with event causation is *a "covering law" model of explanation* according to which some event is explained—or covered—by giving a valid deductive or inductive argument for that event. Such an argument utilizes two features: (1) a universal or statistical law of nature and (2) some initial causal conditions.

A good example of the "covering law" model of explanation is the ideal gas law: PV=nRT (P, V, T are pressure, volume, and temperature, respectively; n is the number of moles of gas; R is a constant). Thus, a covering law model of explanation for the temperature of the gas would look like this:

1. PV=nRT
2. The gas in our container has P_1, V_1, and n_1.
3. Therefore, the gas in our container has T_1.

In this case, we want to explain why our gas has temperature T_1. And we explain this by citing the law of nature (expressed by the ideal gas equation in premise (1), plug in our initial conditions as stated in (2), and we, therefore, explain the fact stated in (3).

Even though many believe the covering law of explanation is a completely adequate standard scientific explanation, others hold—correctly in our view—that while it may be a *necessary* condition of explanation, it is not *sufficient* by itself. It needs to be supplemented by some model that tells us *why* the universal law is true in the first place and why, given a constant volume, a certain amount of gas *must* increase in pressure if the temperature increases. We need a model to explain the necessitation of pressure increase given such and such conditions along with an increase in temperature.

---

[60] Swinburne, R. (2013). *Mind, brain, and free will* (p. 119). Oxford: Oxford University Press.

That model is the ideal gas model. *Gases* are taken to be collections of tiny atoms or molecules that are assumed to be point particles that engage in completely efficient elastic collisions (no loss of momentum). Moreover, *temperature* is reduced to atomic/molecular motion, and *pressure* is reduced to the rate at which the atoms/molecules of the gas collide with a certain area of the container wall. Thus, for example, if we keep the volume constant (as in a pressure cooker) and raise the temperature, the atoms/molecules are going to get agitated and move around much faster (this is what temperature actually is in the ideal gas theory) and this will, in turn, cause more of them to hit the container wall per second (pressure).

So the ideal gas model provides an *explanatory picture* of what is going on, including a *mechanism* (agitating the gas's atoms/molecules) that further explains the ideal gas law. And it provides *necessitation* (nomological or metaphysical) in the explanation: the pressure *must* increase with temperature increase under constant conditions. All the ideal gas law does is to present *contingent correlations* among P, V, n, and T. The law is actually not an explanation at all. Instead, it is the thing to be explained. As it is, it is just a contingent correlation among the relevant factors. The correlation is a mere brute fact without the real explanation (the ideal gas model) that provides necessitation and explanation.

Applying these insights about scientific explanation to our present concerns, it isn't clear how neuroscience could, in principle, undermine dualism. Those who propose neuroscientific anti-dualist arguments do so primarily from advances in mapping psychophysical correlations. Some of the neurological correlates of consciousness were discussed in Chapter 6. Here are a few more. For example, the data from fMRI studies show that brain activity increases in specific areas in conjunction with certain mental acts: the left side of the cerebral cortex when asked to perform language tasks, different but overlapping areas of the left cerebral cortex when listening to someone talk, and a different pattern of cerebral cortex areas when solving mathematical problems.[61] Other fMRI studies show that complex moral reasoning correlates with more intense brain activity. For example, individuals with damage to the lower middle portions of their frontal lobes exhibit an inability to use moral guidelines.[62] But how is any of this a problem for substance dualism? Evans writes,

> Is it a problem that the causal effects should be the product of specific regions of the brain? Why should the fact that the source of the effects are localized regions of the brain, rather than the brain as a whole, be a problem for the dualist? It is hard for me to see why dualism should be thought to entail that the causal dependence of the mind on the brain should only stem from holistic states of the brain rather than more localized happenings.[63]

The correlations between specific mental states and localized brain states is undoubtedly evidence of a relation between mental states and brain states. Although correlates of conscious, the causal direction is not always clear.[64] In cases where mental states do seem dependent on brain states, this needn't phase the dualist. Although neuroscience has helped us understand *how* the mind depends on the brain in some cases, *that* biology plays a role in our thoughts and behavior has been recognized by dualists for centuries. Descartes, for example, held that memory, imagining, and perceiving by the senses were all dependent on the brain.[65]

---

[61] Green, J. D., et al. (2004). The neural bases of cognitive conflict and control in moral judgement. *Neuron*, *44*(2), 389–340.

[62] Bechara, A., Tranel, D., et al. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, (6), 215–225.

[63] Stephan Evans, C. (2005). Separable souls: Dualism, selfhood, and the possibility of life after death. *Christian Scholar's Review*, *34*, 333–334.

The neuroscience objection assumes that if dualism is true, then the mind should not depend on the brain, at least not to the degree that it does. Furthermore, it does not follow from any neuroscientific findings that because the brain is used to do certain things that the brain is what does those things. As Roderick Chisholm observes,

> Many have assumed—quite obviously incorrectly—that from the fact that one thinks by *means of* the brain, it follows logically that it is *the brain* that thinks. We walk by means of our feet, but our feet do not walk in the sense we do (if they did, then they would have feet).[66]

The dualist can always interpret such studies as showing that the nature of the soul is such that while embodied, it is dependent on the brain in a variety of ways. The basic thesis of dualism is open to whatever causal dependence thesis is supported by the evidence. The only way neuroscience could disprove this kind of dualism would be to prove epiphenomenalism. However, as Evans has argued, the discovery that the mind is epiphenomenal would imply that our experience of ourselves as conscious agents is illusory. However, this cannot be the case, as the sciences depend on our self-understanding as conscious agents.[67]

Perhaps the thought is that if dualism is true, it is highly improbable that such neurological correlates would exist. This seems to be what Michael Tooley has in mind where he writes,

> In short, there is a wealth of familiar phenomenon that would be surprising and unexpected, and not at all what one would predict, if the mind were an immaterial substance, but that fall perfectly into place if the mind is, instead, the brain. There is, accordingly, massive evidence against the view that the mind is an immaterial substance.[68]

But what is missing from this argument is an analysis of how the prior probabilities of the psychophysical correlations are improbable given dualism. Let's consider how this might work by setting up physicalism and dualism as rival hypotheses for evaluation against a common body of evidence, the psychophysical correlates, by considering each of their likelihoods given that evidence. Consider the following standard principle of confirmation theory:

---

[64] Richard Swinburne provides compelling arguments that there cannot be any scientific explanation of the neurological correlates of consciousness. The arguments are detailed, but two main points are worth mentioning. First, mental states themselves are not within the domain of science. Second, the laws governing the correlations would be too vast and too complicated or us to have much confidence in the truth of any scientific account. See, Swinburne, R. (1997). *The evolution of the soul* (revised ed., pp. 185–196). Oxford: Clarendon Press.

[65] Descartes, *Meditations on first philosophy*, p. 358, CSM II: 248. See also, *Treatise on man*, 178 CSM I:107, and, *The passions of the soul*, 359, CSM I:343.

[66] Chisholm, R. M. (1991). Mind. In H. Burckhardt & B. Smith (Eds.), *Handbook of metaphysics and ontology* (p. 544). Munich: Philosophia Veril.

[67] Evans, Separable souls: Dualism, selfhood, and the possibility of life after death, p. 335.

[68] Tooley, M. (2009). The appeal to an immaterial mind. In M. Tooley, C. Wolf-Devine, P. E. Devine, & A. Jaggar (Eds.), *Abortion: Three perspectives* (p. 19). Oxford: Oxford University Press.

> *Likelihood Principle*: e counts in favor of h_1 over h_2 if $P(e|h_1) > P(e|h_2)$, where $P(e|h_1)$ and $P(e|h_2)$ represent the conditional probability of e on h_1 and h_2, respectively.

Accordingly, an observation e counts as evidence in favor of hypothesis h_1 over h_2 if the observation is more probable under h_1 than h_2. Let the following values hold: e = the psychophysical correlates; h_1 = dualism; and h_2 = physicalism. On standard accounts, the degree to which e counts in favor of one hypothesis over another is proportional to the degree to which e is more probable under h_1 than h_2. So, how do we determine the degree to which e is more probable given physicalism than it is given dualism?

An evaluation of e alone will not suffice as it tells us nothing about physicalism or dualism. To determine which hypothesis is more probable than the others with respect to e, we must consider physicalism and dualism. But notice that we've left the realm of neuroscience for philosophy. When it comes to analyzing such probabilities, neuroscience itself offers us nothing. As Joshua Rasmussen has argued, it seems reasonable to think that the empirical data regarding psychophysical correlations actually fits better with dualism.[69] Both the location of psychophysical correlates and the neuroplasticity of the brain are precisely what we should expect if dualism is true.

This point has been made by neuroscientists themselves, especially regarding neuroplasticity (the brain's ability to be reprogrammed by the person whose brain it is).[70]

There isn't a purely neuroscientific argument against dualism. We shouldn't take this lightly. Recent studies show that priming individuals with neuroscientific language dissuades belief in dualism, even though the neurological language itself provides no evidence against dualism.[71] Academics are not immune to this effect. We shouldn't be so cavalier with claims as to what neuroscience concludes. There is growing skepticism among neuroscientists regarding exaggerated claims about what neuroscience has or even could show.[72]

---

[69] Rasmussen, J. Against nonreductive physicalism. In *The Blackwell companion to substance dualism* (pp. 336–337).

[70] See, e.g., Beauregard, M. (2007). Mind does really matter: Evidence from neuroimaging studies of emotional self-regulation, psychotherapy and placebo effect. *Progress in Neurobiology*, *81*(4), 218–236; and Schwartz, J., & Begley, S. (2002). *The mind and the brain: Neuroplasticity and the power of mental force*. San Francisco: HarperCollins.

[71] Preston, J. L., Ritter, R., & Hepler, J. (2013). Neuroscience and the soul: Competing explanations for the human experience. *Cognition*, *127*, 31–37.

[72] See, e.g., Shulman, R. (2013). *Brain imaging: What it can (and cannot) tell us about consciousness*. Oxford: Oxford University Press; Satel, S., & Lilienfeld, S. (2013). *Brainwashed: The seductive appeal of mindless neuroscience*. New York, NY: Basic Books; Legrenzi, P., & Umiltà, C. (2011). *Neuromania: On the limits of brain science* (F. Anderson, Trans.). Oxford: Oxford University Press; and Tallis, R. (2011). *Aping mankind: Neuromania, Darwinists and the misrepresentation of humanity*. New York, NY: Routledge.

10.2.7 Neuroscientific Correlations and an Ad Hoc, Bloated Ontology

Neuroscientists are trying to develop laws that govern the relationship between one brain state and another brain state. For example, one proposed law goes like this: “Neurons that fire together, wire together.” In other words, if a group of neurons all fire together, they tend to groove or group together so that they seem wired together to fire simultaneously on a regular basis. So far, so good.

But now suppose that neuroscientists try to go beyond these strictly physical laws, to add new laws that correlate brain states (like C fibers firing) with *mental* states (like feeling pain). We would end up with an unruly list of hundreds of thousands, maybe millions, of brute-fact correlations between various mental and physical states. The expansion of neuroscientific theory to include these correlations would make the theory so utterly complex that the theory itself would be undermined.

As naturalist philosopher Jaegwon Kim notes, if we do not identify mental and physical properties, mental and physical states, and, instead, leave them as two different things that are correlated with each other, then

> all such correlations would have to be taken as “brute” basic laws of the world—“brute” in the sense that they are not further explainable and must be taken to be among the fundamental laws of our total theory of the world. ... But such a theory of the world should strike us as intolerably complex and bloated—the very antithesis of simplicity and elegance we strive for in science.[73]

Because Kim seeks to operate exclusively within a scientific framework, the oddness of such correlations is a problem for him. Thus, mental and physical correlations cannot be explained by neuroscience because they are brute. Not only that, but they are bizarre (they involve entities—irreducible mental entities—that sit oddly and are not at home with the rest of the physicalist ontology). And such correlations are the very antithesis of the physicalist’s favorite epistemic value—simplicity.

But there is another problem—a damaging one for the claim that neuroscience provides explanations for the nature of consciousness and the self—that plagues advocates of neuroscientific adequacy. These correlations are not only odd and brute; they are also contingent. The obtaining of conscious states and the phenomenal properties that are essential to them is contingent relative to their associated “right physical situations.” Our claim is based on the widespread and, in our view, successful employment of a host of thought experiments (e.g., the knowledge argument, the possibility of absent or inverted qualia, zombie worlds, and disembodied existence) that highlight the contingency in question. In regard to the employment of these thought experiments to support the reality and contingency of qualia, Kim observes,

> The [property-dualist] case against qualia supervenience therefore is not conclusive, though quite substantial. Are there, then, considerations in favor of qualia supervenience? It would seem that the only positive considerations are broad metaphysical ones that might very well be accused of begging the question.[74]

---

[73] Kim, J. (2011). *Philosophy of mind* (3rd ed., p. 101). Boulder, CO: Westview Press.

[74] Ibid., pp. 309–310. Brackets ours.

Kim goes on to say that these broad metaphysical considerations amount to the question-begging assumption that physicalism *must* be true.

Second, necessitation allegedly justifies or at least attenuates the adoption of emergent properties in analyzing the obtaining of conscious states by (allegedly) "satisfying" the location-problem: Emergent properties are permitted as long as they are rooted in, nomologically or metaphysically necessitated by, dependent on, and determined by microphysical properties. As David Chalmers correctly notes, without some sort of asymmetric dependency between a strongly emergent property and a physical basal property, "it is not easy to square the view with a contemporary worldview on which everything depends on what is going on in physics."[75] Unfortunately, the contingency of mental/physical correlations (or the very existence of mental states in the first place) violates the staunch naturalist worldview as correctly depicted by Chalmers, and it makes these correlations recalcitrant to neuroscientific explanation. In this way, the correlations weaken the epistemic strength of staunch or faint-hearted naturalism and bloats the naturalist ontology such that these *ad hoc*, brute facts are neither basic, natural, nor plausibly located in the naturalist Grand Story.

10.2.8 The Neuroscientific Parsimony Objection

Some physicalists understand that the neuroscientific data is not altogether in conflict with dualism. For example, Kevin Corcoran and Kevin Sharpe observe,

> Are the deliverances of the neurosciences consistent with the view that human persons are or have an immaterial soul? Yes, they are. But they fit much more seamlessly with a physicalist understanding of human nature and the world.[76]

Corcoran and Sharpe don't clarify what they mean by "fit." Issues of fit are commonly thought of in terms of explanatory simplicity or parsimony. We'll assume that this is what Corcoran and Sharpe have in mind.

On the standard parsimony principle, all things being equal, a more parsimonious explanation of some phenomenon is preferable to a less parsimonious explanation. The following distinction between two kinds of parsimony is standard:

> *Quantitative Parsimony*: A matter of committing to as few individual—token—entities as possible.
>
> *Qualitative Parsimony*: A matter of committing to as few kinds—types—of entity as possible.

Perhaps the claim is that dualism fails the test of *quantitative parsimony* by postulating non-physical properties and immaterial substances. Alternatively, one could argue that dualism fails with respect to *qualitative parsimony* as it postulates non-physical kinds in additions to physical kinds.

---

[75] Chalmers, D. J. (2017). The combination problem for panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (p. 194). Oxford: Oxford University Press.

[76] Corcoran, K., & Sharpe, K. Neuroscience and the human person. In Crisp, Porter, & Ten Elshof, *Neuroscience and the soul* (p. 134).

This argument isn't as straightforward as often presented against dualism. First, qualitative parsimony—counting by kinds of things—is a disputed notion as it is notoriously difficult.[77] Second, suppose that SD is qualitatively less parsimonious than physicalism or naturalism. What follows? This argument gains traction by holding parsimony above other theoretical or explanatory virtues. First, there are, of course, other explanatory virtues, and parsimony is typically treated as among many virtues that a theory or explanation might possess.[78] Second, parsimony is a secondary virtue behind empirical adequacy. That is, parsimony selects from among the list of explanations, with a *ceteris paribus*—all things being equal—clause, after those explanations have been filtered through the relevant data. So, empirical adequacy comes before considerations of parsimony.

The debate between the dualist and the physicalist is, first, one of being able to explain the data. In this book, for example, we've argued that physicalism cannot explain the phenomenal unity data. Parsimony issues are not relevant until we establish that both dualism and physicalism can account for the data. This point is important. The issue is not merely about neural states correlating with states of consciousness. Instead, the issue, at least in terms of phenomenal unity, is one of explaining how the neurological correlates of consciousness all correlate with my, and only my, phenomenally unified consciousness. The neural correlates argument, at least regarding phenomenal unity, may turn out to be significant evidence against physicalism.

Last, it is an ongoing debate if parsimony is a truth-indicator, as the neuroscientific parsimony objection assumes. But there is significant debate that parsimony is nothing more than a pragmatic principle.[79] Indeed, one might agree with Bradley Monton, for example, that we should be skeptical of controversial philosophical conclusions derived from uncontroversial judgments about parsimony.[80] So, it seems that the dualist has various ways of resisting the neuroscientific parsimony argument in particular and neuroscientific objection to dualism in general. This issue is not at all as straightforward as commonly presented.

---

[77] See, Nolan, D. (1997). Quantitative parsimony. *British Journal for the Philosophy of Science*, *48*, 329–334; Huemer, M. (2009). When is parsimony a virtue? *Philosophical Quarterly*, *59*, 216–235; and Jansson, L., & Tallant, J. (2017). Quantitative parsimony: Probably for the better. *British Journal for the Philosophy of Science*, *68*, 781–803.

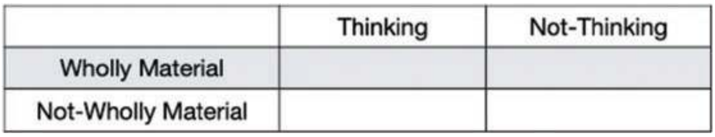
[78] Thomas Kuhn lists several explanatory virtues: e.g., accuracy, consistency, breadth of scope, simplicity, and fruitfulness. See Kuhn, T. (1977). *The essential tension: Selected studies in scientific tradition and change* (pp. 321–322). Chicago, IL: The University of Chicago Press. See Keas, M. (2018). Systematizing the theoretical virtues. *Synthese*, *195*, 2761–2793.

[79] See, e.g., Swinburne, R. (1997). *Simplicity as evidence of truth*. Milwaukee: Marquette University Press; Swinburne, R. (2001). *Epistemic justification*. Oxford: Oxford University Press, chapter 4; Lycan, W. (2002). Explanation and epistemology. In P. Moser (Ed.), *The Oxford handbook of epistemology* (pp. 408–433). New York, NY: Oxford University Press; White, R. (2005). Why favour simplicity? *Analysis*, *65*, 205–210; and Grünbaum, A. (2008). Is simplicity evidence of truth? *American Philosophical Quarterly*, *45*, 179–189.

[80] Monton, B. (2018). The argument from simplicity and the argument from induction: Atheistic induction by Boltzmann Brains. In J. L. Walls & T. Dougherty (Eds.), *Two Dozen (or so) arguments for God: The Plantinga Project* (p. 185). Oxford: Oxford University Press.

10.2.9 Defeating and Turning the Tables on an SD Parsimony Argument

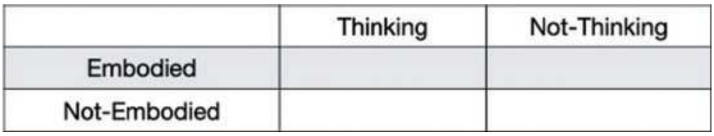
Andrew Bailey has identified a parsimony argument that, *prima facie*, provides support for SD. However, Bailey presents what he takes to be a successful defeater for the argument and, further, turns the argument into one in favor of Monotheistic Physicalism.[81] Bailey begins by granting epistemic parsimony by affirming "that simpler is better," and shows how this principle can be used to defend Monotheistic SD (MSD) over against Monotheistic Physicalism (MP). Keep in mind that his dialectic is among monotheists. Given this limitation, Bailey notes that MSD entails only one kind of (divine or human) thinker—an immaterial one. However, MP postulates two kinds of thinkers—a (divine) immaterial thinker and a human material thinker. Thus, MSD enjoys some support over MP in being simpler. Bailey graphs the situation as shown in Figure 10.1.



	Thinking	Not-Thinking
Wholly Material		
Not-Wholly Material		

FIGURE 10.1 Wholly and Not-Wholly Material.

If we set aside eliminativism and nihilism, and disregard the right column, the MSD claims only one box is filled (the lower-left box) and the MP embraces two boxes in the left column. So MSD can claim that epistemic simplicity is on its side. In response, Bailey invites us to consider another chart, shown in Figure 10.2.



	Thinking	Not-Thinking
Embodied		
Not-Embodied		

FIGURE 10.2 Embodied and Non-Embodied.

Here, both views entail that the left column is filled, and neither can claim the advantage of simplicity. Bailey concludes that since simplicity considerations are sensitive to framing, and unless the dualist provides an argument for preferring the first chart, the dualist simplicity advantage is small (or, we add, non-existent). We want to take up that challenge and offer such an argument here. In a nutshell, we claim that there is an important disanalogy between the two charts that undercuts the analogy between the two charts, and renders (10.2) deeply problematic, and elevates chart 1 as the correct way to depict the dialectic.

Let's focus our attention on the left column in both charts. In (10.1), the two boxes contain members that necessarily, even essentially, belong to their rows or necessarily have no members. The upper box in the column is necessarily empty and the lower box has members—e.g., human persons, angels, God—that essentially belong to that box. This approach to the chart treats the labels for the rows and columns to be subcategories (setting aside worries about the use of solely negative descriptors) and, as we discussed in Chapter 9, a realist construal of categories depicts their members as necessarily or essentially belonging to them (or not). Moreover, given the arguments presented in this and other books on behalf of SD, the upper left box is empty.

---

[81] Bailey, A. M. (2021). *Monotheism and human nature* (pp. 22–24). Cambridge: Cambridge University Press.

However, (10.2) is quite different. The box containing not-embodied thinkers—angels and God for Bailey—does so necessarily. It is part of the very essence of God and angels that they are not-embodied (either angels can contingently be embodied (depending on what embodiment entails) and belong in the upper left box, or they can appear *as if* embodied and belong in the lower left box). However, the upper left box is quite different from the lower left box of (10.2), as well as from its counterpart in (10.1). In (10.2), members belong to the lower left box necessarily and the upper left box contingently.

Regarding the upper left boxes of charts 1 and 2, members who allegedly fit in Wholly Material-Thinking are necessarily members of their row. But that is not the case with (10.2). There, members of that row (Embodied) are contingently in the left column and necessarily in the right one. Thus, the chart itself is a confusing mixture of boxes with contingent and necessary members; the contingency it exhibits violates the very purpose of a realist formulation of categories and is a poor analogy with (10.1).

Bailey would disagree with our argument, most likely because (1) he employs categories with items contingently belonging to them; (2) we beg the question by appealing to dualist arguments to justify the claim that the upper left box of chart 1 is empty. In response to (1), we have already noted in Chapter 8 that, as far as we know, Bailey is the only philosopher in the history of categorial reflection from a realist perspective who adopts a modality of contingency for category membership. He offers no argument for this claim. Until he does, we remain convinced that his analysis of categories is mistaken and confused in its understanding of the purpose of developing a realistic set of categories.

Regarding (2) there is no begging the question in our approach, nor do we know of any other intellectual impropriety in our employment of MSD to justify the claim that the upper left box of chart 1 is empty. We say this because epistemic/explanatory parsimony is a tie-breaker and comes at the end of a debate after the conclusion that the competing views are equal *sans* simplicity considerations. It follows that the appropriateness of appealing to simplicity crucially depends on a fair and accurate presentation of the competing positions. Thus, it is quite appropriate to weigh a full and accurate presentation of MSD and MP *before* appealing to simplicity. And in using SD arguments that entail the emptiness of the relevant box, we have done just that.

10.3 POSTULATING A SUBSTANTIAL SOUL IS EXPLANATORILY IMPOTENT IN EXPLAINING THE OCCURRENCE OF CONSCIOUS STATES

10.3.1 SD and Explanatory Impotence

A frequently raised objection to SD is the claim that postulating a soul does not explain at all the occurrence of conscious states. Thus, Ian Ravenscroft claims that there are "six general features of mental life which a good theory of mental states should be able to explain (or explain away)."[82] Here are the first three features:

---

[82] Ravenscroft, I. (2005). *Philosophy of mind: A beginner's guide* (p. 17). Oxford: Oxford University Press.

1. Some mental states are caused by states in the world.
2. Some mental states cause actions.
3. Some mental states cause other mental states.

Ravenscroft concludes that SD is virtually silent in explaining (1)–(3) with no answer forthcoming. Along similar lines, against the charge that we have no clue as to how consciousness could arise in a physical object, Jaegwon Kim replies: "Is it any easier to understand how thoughts and consciousness can arise in an immaterial substance … Understanding how mentality can arise in something immaterial may be no easier than understanding how it could arise in a material system; in fact, it might turn out to be more difficult."[83] Finally, Eric Olson declares that

> [I]t may well be that we cannot conceive how a material thing could produce thought or consciousness. But then we can no more conceive how an *im*material thing could produce thought or consciousness; no matter how carefully we reflect on the notion of an immaterial substance, we shall never find anything that could account for thought or consciousness.[84]

10.3.2 SD and the Problematic Nature of Soul Stuff

This objection is often raised in conjunction with the notion that a soul is composed of some sort of mind stuff which is virtually impossible to characterize and, if described, it is done so by an inadequate negative one (e.g., immaterial). To cite two examples, Ravenscroft complains that SDs "assert that nonphysical mental stuff is conscious; they do not tell us what it is about nonphysical stuff that facilitates [causes] consciousness."[85] Likewise, Paul Churchland laments that employing "nonphysical" in characterizing a soul "leaves us wanting to ask for something more in the way of a *positive* characterization of this proposed mind-stuff."[86] This objection is ubiquitously employed, and it would have real force if the only way to characterize a soul was by using the *via negative* approach. But as we will see in what follows, that is not what dualists do. Rather, they provide a positive characterization of a soul.[87]

---

[83] Kim, *Philosophy of mind* (3rd ed., p. 46). Physicalists are often so confused about precisely what the issue is that no less a philosopher than Kim conflates the problem of explaining how consciousness could arise in or be caused by an immaterial substance (as expressed in the quote) with the problem of how consciousness could "be possessed by" or "inhere in" a simple, immaterial substance. See Ibid. Clearly, these are *different* problems. As we will see, a failure to distinguish them has lead a number of philosophers to chase an unprofitable red herring.

[84] Olson, E. T. (2007). *What are we?* (p. 153). Oxford: Oxford University Press. See also, McGinn, C. (1999). *The mysterious flame* (pp. 94–95). New York, NY: Basic Books.

[85] Ravenscroft (2005). Brackets ours.

[86] Churchland, P. (2013). *Matter and consciousness* (3rd ed., p. 12). Cambridge, MA: MIT Press.

[87] See, e.g., Moreland, J. P. (2013). A conceptualist argument for substance dualism. *Religious Studies*, *49*, 35–43. For a response, see Bailey, A. (2021). *Monotheism and human nature* (pp. 16–22). Cambridge: Cambridge University Press.

10.3.3 SD Responses to These Objections

These associated objections express a profound and troubling confusion on the part of such physicalists. In short, when used as an explanatory postulate (and we have noted that many SDs are such because they take themselves to be directly aware of their simple souls), SD is not intended to provide a *causal explanation* for the occurrence of conscious states; SD is an ontological theory as to what sort of thing must an entity be to capture adequately several features, not of mental states, but of the thing that *has* them.

In Chapter 1, we listed certain desiderata for a theory of substance and, more specifically, a theory of SD. Here we list an overlapping set that we have tried to satisfy in previous chapters:

1. The appropriate entity for exemplifying mental properties.
2. The appropriate entity for holistically unifying mental properties.
3. The appropriate entity for being an enduring mental continuant.
4. The appropriate entity for being the employer and referent of "I."
5. The appropriate entity for having an irreducible "First-Person Point of View."
6. The appropriate entity for being possibly disembodied as a unified center of consciousness.
7. The appropriate entity for ontologically grounding claims like "Necessarily, thoughts have thinkers."
8. The appropriate entity for exercising active power and teleologically guiding a deliberative process toward an end.

Employing a longstanding, doggedly persistent notion in the history of philosophy—that of a classic substance—the substance dualist refuses to reduce or eliminate these eight features, and instead, uses that notion to explain and ontologically ground them in a fundamental substance. Whether correct or not, the substance dualist holds to a substantial soul because it is the best ontological account of these features, not because a soul best explains what causes conscious states. Thus, *the soul is essentially constituted by actual and dispositional powers of consciousness (and, for some, life) and is not the cause of mental states except in the case of certain agent-causal libertarian acts*. It follows that these complaints about explanatory impotence address a strawman. It also follows that souls are essentially constituted by properties and property-instances, and not stuff. To our knowledge, the only contemporary substance dualist who embraced soul-stuff was Richard Swinburne, a position he has subsequently abandoned.[88]

10.3.4 Responding to a Misguided Physicalist Defeater

Ian Ravenscroft has raised a defeater for our response, or at least for a close cousin to it:

> It might be asserted that the nonphysical mind *just is* conscious, but now the substance dualist must make that claim plausible. In addition, if it's OK to assert

---

[88] Swinburne, R. (1986). *The evolution of the soul* (rev. ed., pp. 153–155). Oxford: Clarendon Press; cf. *Mind, brain, and free will*. (2013). (pp. 33–38). Oxford: Oxford University Press.

> without argument that the nonphysical mind just is phenomenally conscious, why can't the physicalist help herself to the same move? The brain *just is* phenomenally conscious and no explanation of that fact is needed.[89]

There are several problems with Ravenscroft's claims. First, throughout this book, we (and many other substance dualists) have offered a number of arguments for why there is a soul to ground adequately or account for a number of features listed in Chapter 1 and Figures 10.1 and 10.2. So we (and most substance dualists) do not simply assert that the soul just is conscious without supporting arguments. Second, we have given reasons throughout this book for the claim that matter construed as a mass of stuff or an aggregate/system of separable parts fails to provide that adequate ground/account. Nor are we unique among substance dualists in this regard. Whether or not the arguments involved in our first two rejoinders are successful, their presence is enough to refute Ravenscroft's charges.

Finally, Ravenscroft's complaint against the substance dualist's (alleged) assertion "that the nonphysical mind *just is* conscious" is wide of the mark. Why? Using "soul" rather than "mind" for consistency, it is just false to claim that there is no positive characterization of a soul. Dualists hold that it is a substance essentially characterized by the actual and dispositional properties of consciousness (and, perhaps, life). Some would add the categorial feature of being a substance. Ravenscroft is misled by his unfortunate view that substance dualists take the soul to be a mass of negatively characterized stuff that *causes* conscious states. In fact, the substance dualist characterizes the soul as essentially *that which* ______, where the blank is filled with (at least) the eight features above (e.g., exemplifies mental properties, holistically unifies mental properties).

To unpack this last claim in more detail, let us begin with an analysis of the notion of a substance. For present purposes, let us grant that a substance is a primitive, underived unity of actual and potential properties or inseparable parts at a time, it sustains absolute sameness through accidental change, and so on regarding the list above, and it has an essence that answers the most fundamental question "What kind of thing is this?" and that grounds its membership in its *infimae species*.

Note that the characterization just given is a formal one. It merely states necessary conditions for anything to count as a substance without giving material content necessary to distinguish one kind of substance from another. When it comes to characterizing specific kinds of substances, it seems that there are only two things one can do: state the formal criteria and list a set of dispositions or potentialities as the material content for a specific kind of substance. For example, a characterization of an atom of gold as a substance might go something like this: It is a primitive, underived unity of actual and potential properties or parts at a time (having the power to melt at such and such degrees, dissolving in *aqua regia*, having atomic number 79 which, in turn, is the potential to attract a certain number of electrons, to resist inertial changes to a degree thus and so), it sustains absolute sameness through accidental change, and it has an essence (being gold) that answers the most fundamental question “What kind of thing is this?” and that grounds its membership in its *infimae species* (the natural kind *gold*).

---

[89] Ravenscroft, I. (2019). Why reject substance dualism? In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 267–282). Oxford: Wiley-Blackwell. The quote is on 272.

Note that the formal aspects of the definition of gold are the same as those for a substance in general, and the material content is listed in parentheses. Note also that the various properties listed in parentheses that give material content to “being gold” so as to distinguish it from “being iron” are dispositions or potentialities, properties that are actualized by a gold atom if certain conditions obtain (e.g., gold will dissolve in *aqua regia* if placed in it).

Substances are more than their potentialities; indeed, it seems reasonable to say that something can have potential properties only if it has actual properties: No potentiality without actuality. Unfortunately, it is very difficult to say what the essence of a substance is without referring to its potentialities. Thus, we say that gold is *that which* has the potential to do x, y, and z; iron is *that which* has the potential to do f, g, and h; and so on.

By way of application, it is hard, if not impossible, to characterize adequately a specific kind of substance without reference to the various potentialities it has by its very nature and which distinguish its essence from the essences of other kinds of substances. That’s just the way it seems to be with substances. Still, it is important to realize that the dispositional nature of the material content of a substance is true of physical substances just as much as spiritual substances. The problem here resides in the limitations of characterizing substances in general and is not uniquely troublesome for souls.

10.3.5 A Response to Bailey’s Magical Mystery Tour

Andrew Bailey cites several philosophers who either assert or provide a thought experiment—Leibniz’s famous mill thought experiment being the paradigm case—for the thesis that it is simply impossible that a material object has phenomenal consciousness since, upon examination, it is quite inexplicable how such an object could cause or in some other way bring about such consciousness.[90] There can be no full or even partial explanation of this fact. No mechanism, no description at all, brings us any closer to understanding how this is done. Let’s call this the *Impossibility Argument* (IA).[91]

Bailey provides a thoughtful response to IA. He begins by saying that a phenomenon is magic to the degree that it is modally and explanatorily independent of the material world. Thus, magic is degreed. It follows that all facts about God are purely magic, e.g., his thoughts and sensations. Since the monotheist is committed to pure magic regarding God’s thinking and other mental states, there is just no “how” to God’s thinking. The mysterious view is the notion that there is no “how” to the act of thinking (having phenomenally conscious states); thus, thinking is a mystery.

Bailey concludes that the IA does not show that it is impossible for a material object to think. Instead, it shows that thinking itself is a mystery. And there is parity here regarding immaterial and material objects, human or divine, so the IA fails.

But Bailey is not content with a draw. He goes on to argue that if we examine this issue further, we discover that the physicalist has an advantage over the SD. Bailey observes that SDs acknowledge a dependency, even causal, relation between the body and its parts and the obtaining of phenomenal conscious states. Thus, SDs admit that phenomenal consciousness is "tethered" to a body in that it is at least partly generated by the physical body. But this acknowledgment provides an advantage for subject physicalists over SDs. The physicalist can embrace *Tethered Body*: When a human person (who is a wholly material being) thinks, it is at least partly because her parts are in a certain physical state. But the SD is stuck with *Tethered Soul*: When an embodied human person (who is an immaterial soul) thinks, it is at least partly because the parts of her body are in a certain state.

---

[90] Bailey, pp. 24–32.

[91] See Chapter 6 for our versions of mill-like arguments from the phenomenal unity of consciousness.

SDs are saddled with the view that one's thinking is partly due to something disjoint from that person, so one's thinking is partly generated by "leaping across the divide from body to soul," and this is an additional mystery the *Tethered Body* position does not have to face.

This is an important argument, and SDs need a reply to it as we now seek to provide. We see at least three problems with Bailey's argument.

First, we think Bailey has misunderstood the nature of the argument presented by the mill experiment and others he cites. Since we don't want our reply to involve exegesis, regardless of any such misunderstanding, we believe Bailey's critique does not touch the point SDs *ought* to make with the quotes and examples cited by Bailey for IA. Recall from earlier in this chapter that we distinguished two interpretations of the question, "How can mental entities causally interact with physical entities?" The *Mechanistic Question* is a request for a mechanism that intervenes between the mental and physical, a description of which explains how it is that such interaction takes place. It is clear that some form of the *Mechanistic Question* is the one Bailey attributes to those promoting the Impossibility Argument and, accordingly, the one on which he focuses.

However, we think the *Skeptical Question*—an expression of skeptical exasperation, perhaps based on the differences between mental and physical entities—is the one behind or ought to be behind IA, and it expresses what advocates take to be two forms of insufferable difficulties for anyone who believes that material objects can think. (1) The origin of consciousness problem. How can purely physical entities that engage in purely physical processes give rise to something brand new and completely non-physical, all at one go? Inspect the mill all you want. You will not find even dispositions for consciousness. We developed this problem in Chapter 2 in our discussion of emergence and will not repeat those ideas here.

(2) The inadequacy of a scattered object or mereological aggregate to serve as a proper candidate for being a subject of consciousness. This problem resembles one raised by Peter Unger, though we want to say more than Unger did.[92] Above, we listed eight desiderata for a theory of substance and a theory of a conscious substantial subject. Throughout this book we have argued for the superiority of an SD account of them. Here, the point is this: Given these eight items that characterize a conscious substantial subject, the different statements of the *Skeptical Question* raise the point that wholly material objects just aren't the right sorts of things to satisfy these desiderata.

Whether or not there are adequate physicalist answers to these skeptical questions, it is still the case that Bailey's focus on the *Mechanical Question* is wrong-headed.

Second, Bailey is misguided when he says that the monotheist is committed to pure magic and mystery regarding God's thinking since there just is no "how" to God's thinking. We say this for two reasons. For one thing, granting the legitimacy of Bailey's understanding of the "how" question, from the fact that God's thinking is pure magic in that it is modally and explanatorily *independent of the material world*, it does not follow that there is no "how" to God's thinking. That "how" could be spelled out satisfactorily even if it did not appeal to factors depending on the material world. One of us has provided a broad-stroke account of how God becomes aware of things and how this relates to God's propositional knowledge.[93] Successful or not, the mere possibility of such an account is sufficient to show the above entailment is wrong. Thinking is a mystery in Bailey's limited sense of that term, but it need not be such in a broader, typical understanding of mystery that rules out any answer to the "how" question.

---

[92] Unger, P. (2004). The mental problems of the many. In D. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. I, pp. 196–222). Oxford: Clarendon Press.

For another thing, answering the *Mechanistic Question* plays little or no role in formulating or defending SD. Still, it plays an essential role in providing a good explanation of the behavior and modal abilities/effects of complex, wholly material objects (setting aside basic fundamental issues such as asking why an electron is negatively charged). We developed these issues more fully in Part III above and will not rehearse them here. Suffice it to say that answering the *Mechanistic Question* is unimportant for SD but crucial for physicalists who claim we are complex, wholly material objects. And if someone replies that such an object simply generates thought as a basic, primitive fact, then one faces the origin of consciousness problem. It should be said that it is this problem that has motivated much of the present return to various forms of panpsychism, so the problem is widely acknowledged to be a serious one.

Third, it is hard to take Bailey's complaint about Tethered Soul relative to Tethered Body seriously, especially his characterization of the alleged dualist commitment to the idea that one's thinking is partly generated by "leaping across the divide from body to soul." To begin with, does anyone really think that, according to dualists, thinking is generated by "leaping across" some divide? The notion of "leaping across" conjures up images of some big gap that something must jump over to get to the other side. The vast majority of SDs actually hold that the soul and body are, in one way or another, directly in contact with no gap separating them.

To be sure, there is a boundary between the two (for Cartesian dualists), but causal interaction is direct and immediate, not in virtue of some intermediate entity standing in the gap. And different versions of staunch hylomorphism make the integration between soul and body deeper and more intimate than a mere causal link. Indeed, many in this camp do not hold the body to be entirely physical, claiming that a body includes the constituting holenmeric presence of its soul. No gap here.

We see no reason why causal interaction between soul and body isn't sufficient to respond to Bailey's complaint here. Bailey accepts such interaction between God and the world, and there is no reason why an appeal to such interaction regarding human persons is not fully adequate to defeat the claim that *Tethered Body* is preferable to *Tethered Soul* unless Bailey accepts a general rule that we should prefer immanent over transient causation. We doubt that he does, and, in any case, we reject such a rule. We conclude that Bailey's arguments against IA are misguided and ineffective.

---

[93] Moreland, J. P. (2021). *A simple guide to experience miracles* (pp. 53–56). Grand Rapids, MI: Zondervan. The account presupposes that God entered time (construed according to the A-series view) at the beginning of creation. It also assumes that concepts and propositions are mental properties that God's mind exemplifies and are the immanent content of God's thoughts.

PART VI

CHARTING A FUTURE FOR SUBSTANCE DUALISM

CHAPTER 11

New Research Programs for 21st Century Substance Dualism

Indeed, substance dualists have focused primarily on defending substance dualism and critiquing alternative views rather than developing detailed SD accounts or models of specific phenomena. Some take this to be a damning strike against SD. Thus, Ian Ravenscroft says,

> What's striking about SD is that it is a theory in only the most attenuated sense of the word. Substance dualists have nothing to say about the components of the nonphysical mind or how those components are structured into a working whole. SD has no account of reasoning, perception, or coordination of action. It is silent on mental illness, memory, emotions, dreams, and imagination. SD offers nothing on psychological development in children or cognitive decline in the elderly.[1]

Ravenscroft's claim is overblown and contestable, and we responded to it in considerable detail in Chapter 10. Still, the underlying point stands. The next phase of SD must include numerous robust research programs. Although some have raised this problem, it deserves detailed and sustained attention.[2]

---

[1] Ravenscroft, I. (2018). Why reject substance dualism? In J. J. Loose, A. J. L. Menuge, & J.P. Moreland (Eds.), *The Blackwell companion to substance dualism* (p. 280). Oxford: Wiley-Blackwell.

[2] Stewart Goetz and Charles Taliaferro suggest four areas of inquiry: (i) ultimate background issues of the naturalism–theism debate; (ii) the concept of what is physical; (iii) cross-cultural inquiry; and (iv) value inquiry. Goetz, S., & Taliaferro, C. (2011). *A brief history of the soul*. Oxford: Wiley-Blackwell, Ch. 8.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

We divide this chapter into four broad research projects—conceptual, methodological, developmental, and applied—within which we make proposals, some more detailed than others. We intend to highlight the fact that SD has a robust research program worth pursuing and suggest particular research projects we think are likely to bear fruit. We also hope these proposals inspire fellow substance dualists to join us in further developing these ventures.

11.1 CONCEPTUAL RESEARCH PROJECTS

11.1.1 Charting the Logical Space of Mere Substance Dualism

Identifying the logical space of SD will reveal numerous, novel, and likely rewarding ways to develop SD.[3] Yet, the logical space of SD is as unexplored as it is acknowledged.[4] SD, it is popularly assumed, is committed to what many would call the four *Cartesian Commitments*:[5]

(i) *Non-spatial*: Souls are not spatially located or extended.
(ii) *Simplicity*: Souls have no substantial parts but are simple substances.
(iii) *Independence*: Souls are in no way dependent upon the physical world for their coming into existence, persistence, development, or powers of consciousness and cognition.
(iv) *Interaction*: The soul and body of a human person causally interact with each other.

It is not difficult to show that SD does not entail (i)–(iv) and that surprisingly the usual SD suspects—Plato, Augustine, and Descartes—did not accept all *Cartesian Commitments*. We will focus on (i) and (iii).

11.1.1.1 The Historical Argument

Plato, Augustine, and Descartes rejected *Independence*. Like Plato,[6] Augustine held that the soul imparts life to the body of all living things.[7] Descartes explicitly rejected the notion that the soul resides in the body "as a pilot resides in a ship" but held that the soul and body form a natural unity "most closely joined" and "as if intermixed,"[8] referring to the soul as a "substantial form," "substantially united" with the body.[9] Although Plato regarded the normative status or flourishing of the soul as radically independent from the body,[10] Augustine and Descartes recognized the soul's dependence on the body for its flourishing.[11] Last, all three held that at least some forms of cognition are bodily processes. For Descartes, memory, imagining, and perceiving via the senses depend on the brain.[12]

---

[3] We are aware of only one detailed, albeit incomplete, treatment of the logical space of SD: Meixner, U. (2004). *The two sides of being: A reassessment of psycho-physical dualism*. Paderborn: Mentis-Verlag, chapter 1.

[4] For exceptions, see Goetz and Taliaferro, *A brief history of the soul*; and Eberl, J. T. (2010). Varieties of dualism: Aquinas and Swinburne. *International Philosophical Quarterly, 50*(1), 39–56.

[5] See, e.g., Kim, J. (2011). *Philosophy of mind* (3rd ed.). Boulder, CO: Westview Press; Churchland, P. (2013). *Matter and consciousness* (3rd ed.). Cambridge, MA: MIT Press; Searle, J. (2004). *Mind: A brief introduction*. Oxford: Oxford University Press; Heil, J. (2012). *Philosophy of mind: A contemporary introduction* (3rd ed.). New York, NY: Routledge.

[6] Plato, *Phaedo*, p. 105c–d.

[7] On animal souls, see Augustine. (1993). *The city of God* (M. Dods, Trans., p. VII.29). New York: The Modern Library; Augustine, *The Trinity*, X.2.6. On human souls, see, *The magnitude of the soul*, 13.22; cf. 33.72. On plant souls, see Augustine. (1947). *Immortality of the soul*. In L. Schopp (Trans.), *The fathers of the church* (Vol. 4). Washington, DC: Catholic University of America Press, 3.3; and Augustine, *The magnitude of the soul*, 33.70.

SD and *Independence*, even in the distinction between mind/soul and body, are not as radical or uniform as popularly portrayed. As Rebecca Copenhaver observes,

> Dualists of the early modern age did not propose two worlds: a manifest world of material things, and a second, ghostly animation, invisible and submerged. If that is Cartesian dualism, no one in the early modern period proposed or promoted it, including Descartes.[13]

Even the SD views of Late Antiquity are not as uniform as is assumed.[14] That Plato championed a pre-Cartesian version of SD is easily contested.[15]

Last, regarding the *non-spatial* thesis, Plato and Descartes held that the soul is not spatially extended.[16] However, both Augustine and Descartes contended that *the soul is located but not extended in space*, wholly present in each space that its body is present.[17] As we will see shortly, contemporary defenders of SD also reject at least one *Cartesian Commitment*.

---

[8] René Descartes, *Meditations on first philosophy*, Sixth meditation, *SCM* I, 81. See also *Discourse on method*, *CSM* I, 141, and *Objections and replies*, *CSM* II, 160. Unless noted otherwise, we will use the standard references for the three-volume set, *Descartes selected philosophical writings*: *CSM* I*, CSM* II*, and CSMK* III.

[9] René Descartes, *Letter to Regius*, *CSM* III, 207–208; and 243, respectively. Descartes' view seems closer to scholastic-Aristotelian theories than is popularly portrayed. See, e.g., Hoffman, P. (1986). The unity of Descartes' man. *The Philosophical Review*, *95*, 339–369; and Skirry, J. (2005). *Descartes, and the metaphysics of human nature*. London and New York: Thoemmes-Continuum Press.

[10] Plato, *Phaedrus*, p. 250c.

[11] See, e.g., Augustine, *The magnitude of the soul*, p. 36.81. Descartes, *Meditations on first philosophy*, *CSM* II, 56.

[12] See, e.g., Descartes, *Treatise on man*, *CSM* I:107; Descartes, *Objections and replies*, *CSM* II, 248; Descartes, *Discourse on the method*, *CSM* I, 127; and Descartes, *The passions of the soul*, *CSM* I, 341, 343.

[13] Copenhaver, R. (2019) Introduction to Volume 4. In R. Copenhaver (Ed.), *Philosophy of mind in the early modern and modern ages: The history of the philosophy of mind* (Vol. 4, p. 3). New York, NY: Routledge.

[14] For a detailed exploration of competing SD views in Late Antiquity, see, Marmodoro, A., & Cartwright, S. (Eds.). (2018). *A history of mind and body in late antiquity*. Cambridge: Cambridge University Press.

[15] Marmodoro, A., & Cartwright, S. (2018). Introduction. In A. Marmodoro & S. Cartwright (Eds.), *A history of mind and body in late antiquity* (p. 3). Cambridge: Cambridge University Press.

[16] Descartes, *Meditations* VI, 190.

[17] Augustine, *The Immortality of the Soul*, 16.25; and Augustine, *The trinity*, VI.2.8. Regarding Descartes, see, Rozemond, M. (2003). Descartes, mind-body union, and holenmerism. *Philosophical Topics*, *31*, 343–367.

11.1.1.2 The Substance Argument

Exploring possible models of substance is a further way to chart the logical space of SD.[18] Given the various understandings of substance, there will be at least just as many types of substance dualism. Here is an incomplete list:

> *Aristotelian Primary Substance*: x is a substance $=_{\text{def.}}$ x is that which is neither said of a subject nor in a subject.
>
> *Aristotelian Hylomorphic Substance*: x is a substance $=_{\text{def.}}$ x is composed of matter and form.
>
> *Substratum Substance*: x is a substance $=_{\text{def.}}$ x is a substratum, a propertyless or bare particular that unifies the properties which inhere in it.
>
> *Inherence Substratum Substance*: x is a substance $=_{\text{def.}}$ x is a complex or collection of a substratum together with the properties which subsist or inhere in the substratum.
>
> *Independence Substance*: x a substance $=_{\text{def.}}$ x can exist without any other entity existing.
>
> *Collectionist Substance*: x is a substance $=_{\text{def.}}$ x is a collection or set of either properties or of concrete entities (e.g., tropes, sense-data, events).
>
> *Powers Substance*: x is a substance $=_{\text{def.}}$ the powers of x are not reducible to the powers of any entity y.
>
> *Substance Pluralism*: x is a substance $=_{\text{def.}}$ x is either a y or a z or … under specified condition(s) C (e.g., the soul is a *Substratum Substance*, and the body is a *Collectionist Substance*).

These distinctions are important. For example, there will likely be new and varied arguments for SD based on distinct notions of substance and perhaps from a general substance dualism.[19] Additionally, objections to SD may succeed or fail one theory of substance or another. Standard interaction problems do so on the supposed difficulty of causal relations between substances understood in a particular way and fail to consider the available SD options seriously. Recognizing this logical space provides ways to advance the discussion.

11.1.1.3 Varieties of Contemporary Substance Dualism

Almost no contemporary substance dualist holds to the complete set of Cartesian Commitments.

Consider three versions of SD: Cartesian, Emergent, and Staunch Hylomorphic Substance Dualism. We will explore the details of these in the embodiment section below. For now, we merely note the existing SD diversity. Here is the first view.

> *Cartesian Dualism*: Mere SD is true; the human mind has only mental properties, the body has only physical properties, and embodiment is a mere causal relation.

---

[18] The follow is informed by the treatment of Joshua Hoffman and Gary Rosenkrantz. See, Hoffman, J., & Rosenkrantz, G. S. (1994). *Substance among other categories*. Cambridge: Cambridge University Press, Ch. 2 and 3; and Hoffman, J., & Rosenkrantz, G. S. (1997). *Substance: It's nature and existence*. New York, NY: Routledge, Ch. 1.

[19] For example, Graham Clay suggests that one might attempt to argue that Locke should be understood as a mind–body substance dualist from, what he argues is, Locke's general substance dualism of material and immaterial substances. See his, "The Contours of Locke's General Substance Dualism," *Journal of Modern Philosophy* (forthcoming).

Some self-described Cartesians (e.g., Stewart Goetz and Charles Taliaferro) reject the mere causal relation view of embodiment as reductionistic, defending what they call

> *Integrative Substance Dualism*: Cartesian Dualism—minus the causal embodiment thesis—is true, and embodiment is a complex of various intimate relations the soul/person has with his/her body (e.g., epistemic, structural, and effective).[20]

Here is another SD view.

> *Emergent Substance Dualism*: Much of Cartesian Dualism is true; however, the soul naturally emerges from an organism's nervous system, on which the soul's existence and development ordinarily depend;[21] it is only simple by lacking substantial parts; it is spatially located, extended, and divisible into separate substances (e.g., split-brain conditions).[22]

Last, consider a specific kind of hylomorphic SD view.[23]

> *Staunch Hylomorphic Substance Dualism*: Mere SD is true; the essence of the soul determines the soul's natural kind and contains, as a primitive unity, powers to act as the first internal efficient cause of (immanent causation) and teleological guide for the internal structure and development of the body; which is a complex psycho-physical mode of the soul.[24]

Other forms of SD in the literature include E. J. Lowe's non-Cartesian SD, Howard Robinson's idealism, Robin Collins's dual aspect SD, and the bodily soul view of C. Stephen Evans and Brandon Rickabaugh, to name a few.[25] Not only do these SD views reject one or more Cartesian Commitments, they add substantive philosophical theses, displaying not only part of the SD logical space, but that these spaces have been occupied for some time, yet ignored by SD detractors.

---

[20] Taliaferro, C. (1994). *Consciousness and the mind of God* (pp. 114–122). Oxford: Oxford University Press.

[21] Hasker, W. (1999). *The emergent self* (pp. 189–190). Ithaca NY: Cornell University Press; Hasker, On behalf of emergent dualism, p. 79; and Hasker, The dialect of soul and body, p. 216.

[22] See, e.g., the following by Hasker, W. (2012). Is materialism equivalent to dualism? In B. P. Göcke (Eds.), *After physicalism* (p. 196). Notre Dame, IN: University of Notre Dame Press; Do my quarks enjoy Beethoven? (2016). In T. M. Crisp, S. L. Porter, & G. T. Elshof (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology*. Grand Rapids, MI: William B. Eerdmans Publishing Company; and Persons and the unity of consciousness. (2010). In R. Koons & G. Bealer (Eds.), *The waning of materialism* (pp. 175–190). Oxford: Oxford University Press.

[23] Robert Koons describes seven contemporary versions of *Staunch Hylomorphism* (see Chapter 8), which, for our purposes, we may call at least some of them *Staunch Hylomorphic Substance Dualism*. See Koons, R. (2014). Staunch vs. fainthearted hylomorphism: Toward an Aristotelian account of composition. *Res Philosophica*, *91*(2), 151–177.

[24] We have acknowledged in earlier chapters that there are several versions of hylomorphism. See, e.g., Moreland, J. P. (2018). In defense of a Thomistic-like dualism. In J. Loose, A. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to mind-body dualism* (pp. 102–122). UK: Blackwell; Voss, S. (1992). Understanding eternal life. *Faith and Philosophy*, *9*(1), 3–22; Barnes, G. (2004). Is dualism religiously and morally pernicious? *American Catholic Philosophical Quarterly*, *78*(1), 99–106; Barnes, G. (2002). Should property-dualists be substance-hylomorphists? *Proceedings of the American Catholic Philosophical Association*, *75*, 285–299.

[25] See, e.g., Lowe, E. J. (2018). Non-Cartesian substance dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 168–182). Oxford: Wiley-Blackwell; Robinson, H. (2016). *From the knowledge argument to mental substance: Resurrecting the mind*. Cambridge: Cambridge University Press; and Collins, R. (2011). A scientific case for the soul. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis: Investigation into the existence of the soul* (pp. 222–246). New York, NY: Continuum.

After stating various distinctions among early modern substance dualists, Copenhaver concludes, "There are more dualisms than are dreamt of in the philosophies of the current curricula in Anglophone universities."[26] The logical space of SD is vast and ripe for discovery.

11.2 METHODOLOGICAL RESEARCH PROJECTS

11.2.1 Empirically Informed and Testable SD

One methodological approach is to develop various empirical tests for SD. A version of this project was ventured by a John Templeton Foundation–funded team trained in philosophy, physics, neuroscience, psychology, and linguistics. One result is a fascinating book that treats SD as a testable hypothesis. Accordingly, SD can function as a hypothesis, "an integral part of a complex overall theory of the true nature of human beings."[27] The research team approached SD as a hypothesis that must be connected to non-eliminable, observable empirical data that would make unanticipated true predictions, subject to thorough testing, refinement, and the development and analysis of alternatives. Our point here is twofold. Contemporary substance dualists take the importance of SD being empirically informed and testable, and more substance dualists should take up these projects.

On the flip side of this suggestion, we recommend that SDists do more work in providing detailed critiques of physicalist claims according to which physicalism gains considerable support *vis-a-vis* dualism by being explanatorily superior in explaining and predicting empirical phenomena. Paul Churchland goes so far as to claim that certain predictions about chimerical colors not only demonstrate the superior explanatory power of physicalism regarding empirical data, but also that the confirmation of certain explanatory predictions provide strong support for reductive (type-identity) physicalism.[28] These kinds of assertions need more responses than dualists have been providing.

11.2.2 Ontologically Serious Philosophy of Mind

Chapter 2 discussed the ubiquitous rejection of First Philosophy and the adoption of the Standard Picture as the correct way to do philosophy, especially philosophy of mind. We see no good reason to accept these positions. In fact, they place artificial restrictions on what counts as acceptable work in the field. Further, they lead to the non-parsimonious multiplication of *ad hoc* brute facts and the ridiculous over-employment of just-so evolutionary stories as a substitute for legitimate argumentation. A classic example of the former is the hackneyed utilization of emergent properties to supposedly solve all kinds of problems. Speaking of those who do this, Mariusz Tabaczek rightly bemoans the fact that

---

[26] Copenhaver, 4.

[27] Baker, M. C., & Goetz, S. (2011). Introduction. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis: Investigation into the existence of the soul* (p. 14). New York, NY: Continuum.

[28] Churchland, P. (2005). Chimerical colors: Some phenomenological predictions from cognitive neuroscience. *Philosophical Psychology, 18*(October), 527–560.

> "what appears to them as a consistent, well-thought-out, and coherent theory [emergentism], may seem in the eyes of an Aristotelian metaphysician to be nothing but a confusing welter of ideas, characterized by a maddening ambiguity of terms."[29]

A well-known expression of frustration with the latter is an insight proffered by Thomas Nagel. Speaking of the naturalist fear that there may be a God—what he calls the "cosmic authority problem"—Nagel observes:

> My guess is that this cosmic authority problem is not a rare condition and is responsible for much of the scientism and reduction of our time. One of the tendencies it supports is the ludicrous overuse of evolutionary biology to explain everything about life, including everything about the human mind.[30]

As we see it, it is time for substance dualists to lay aside these shackles, along with naturalism and scientism, that artificially constrain their ability to approach the field in a distinctively philosophical way. As we have tried to show throughout this book, the best way to clarify and address problems in the philosophy of mind is to approach the discipline with a rich and rigorous set of metaphysical categories and ontological distinctions. And we recommend that substance dualists join together to do research projects in philosophy of mind by focusing more on issues in analytic ontology. When we write, we should lay out our metaphysical framework (as we do in Chapter 3) before launching into specific topics.

11.2.2.1 Properties, Powers, and Substances

If SD is correct, we think a realist adoption of properties, classic substances, and essences is the best way to approach issues regarding consciousness and the self.[31] Given a substance ontology, a causal-powers view of causation is either entailed or at least the most plausible view. Accordingly, we recommend that substance dualists embrace causal powers. Further, in the spirit of Husserl and his contemporary followers, substance dualists need to develop views of acts of knowing that best comport with substance dualism. In our view, such an ontological framework of properties, substances, essences, and causal powers is much more appropriate for developing an ontology of human persons than is the currently dominant approach that begins with separable parts and derivative wholes and uses extensionalist mereology, composition, or constitution to approach the topic of what we are. Suppose we are substantial souls and the fundamental axiom of being a substance is true (no substance has substantial parts). In that case, this dominant approach, often funded by microphysical priority and some form of the atomic theory of matter, is a non-starter.

---

[29] Tabaczek, M. (2019). *Emergence* (p. 63). Notre Dame, IN: University of Notre Dame Press.

[30] Nagel, T. (1997). *The last word* (p. 131). New York, NY: Oxford University Press.

[31] For a recent example by a substance dualist, see, Weir, R. S. (2021). Bring back substances! *The Review of Metaphysics*, *75*(2)(298), 265–308; and Weir, R. S. (work in preparation). *Metaphysics and the mind-body problem: Why the twenty-first century still needs nonphysical substances.*

11.2.2.2 Parts, Wholes, and Unity

Does our claim in the previous section rule out the importance of part/whole issues for doing philosophy of mind? Not necessarily. We think that the employment of inseparable parts (aka modes), ontologically prior substantial wholes, and internal relations are crucial for work in philosophy of mind. Such a framework provides an adequate ontological analysis of the sort of unity that substances, especially souls, have. It also allows one to clarify the soul's simplicity without denying that several intrinsic entities (e.g., faculties) are among the soul's constituents. We recommend that substance dualists do more research on the possible benefits of some version of staunch hylomorphism regarding the human person, especially the body. In our view, the ontological framework that we claim follows most naturally from substance dualism is far superior to adopting a dynamic-systems approach that gratuitously slaps internal relations and inseparable parts onto a complex system of parts and processes.

11.2.2.3 The Return of Teleology

With the rise of the Mechanical Philosophy and the associated revival of the atomic theory of matter in the 17th century, irreducible teleology vanished over the next three centuries. Boyle and the chemists removed teleology from chemical change. Newton did the same regarding mechanics, Darwin with biology, and Freud and Skinner completed the project in psychology and the study of human action and reasons-explanations. By "the Mechanical Philosophy," we do not include the notion that the only causation is efficient causality by contact. To be sure, those like Descartes did embrace this position. But as John Yolton has reminded us, in the early days of atomism and the Mechanical Philosophy, inert corpuscularianism eventually gave way to dynamic corpuscularianism.[32] According to the former, a corpuscle (atom) of matter was dead, inert, and passive. It could be set in motion only by contact with another moving corpuscle. All causation was efficient causality by contact. According to the latter, forces of attraction were attributed to the inner nature of corpuscles. The result was that while causality remained limited to efficient causality, contact was no longer a necessary condition for such causality to obtain. Yet both views were expressions of the Mechanical Philosophy. We mean by the Mechanical Philosophy an approach to macro-objects, including animal bodies, according to which they are machines that engage only in efficient causality and are mereological aggregates of separable parts standing in external relations to each other.

Interestingly, in the last few decades, irreducible teleology has been making a comeback, especially in psychology, biology, and the philosophy of action, in which philosophers use teleology in analyzing reasons-explanations. In psychology, a 2013 article sounded a clarion call for psychology to recover teleology.[33] And the growing emergence of organicists in biology who embrace teleology is a part of this comeback.[34]

---

[32] Yolton, J. (1983). *Thinking matter: Materialism in eighteenth-century Britain* (pp. 3–152). Minneapolis: University of Minnesota Press.

[33] Seligman, M., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science, 8*, 119–141.

[34] See Chapter 7, notes 105–107.

In light of these developments, we would like to see substance dualists engage in future research into the nature of the soul, the body, and agency theory in an attempt to see how fruitful teleology may be in solving problems and opening up new areas of investigation.

11.2.3 Preserving the Person/Subject

More than anyone else, the substance dualist has substantive reasons to preserve the person, to take the person as a fundamental unit of analysis and explanation. This means an overall rebellion against sub-personal, sub-rational, and mechanistic explanations when at all possible. Like panpsychism, which frontloads phenomenal consciousness into the fundamental level, SD does so with subjects of consciousness or persons. So whatever goods come about from the former comes about for the latter and many more. We should recall a distinction mentioned in Chapter 7 between an analysis and an assay of some entity, a person in this case. In our usage, an analysis of a person is reductive. In one way or another, it reduces "person" to a set of non-personal bodily or psychological states, thereby eliminating persons from one's ontology. By contrast, an assay is an inventory of some entity—a person—that tells us what a person is, namely, such and such constituents. An assay is not reductive. For example, one could hold that a person is identical to personhood being exemplified by a primitive this. An assay retains persons in one's ontology and clarifies what persons are.

At least two methodological implications follow. First, the substance dualist can approach the unity of consciousness and the binding problems in a unique SD approach by developing a phenomenology that does not betray the ubiquitous direct acquaintance of oneself, revealing the SD features. Second, the substance dualist has a project of exploring aspects of reality that are fundamentally personal and what this might entail.

11.2.4 Preserving the First-Person

In contemporary philosophy of mind, it is virtually axiomatic that one should dismiss the "first-person point of view" and the unique information gained thereby. And for those who still retain it in their methodology, epistemology, or ontology (inclusive disjunction), it is usually mischaracterized. An example of the first approach is Paul Churchland's dismissal on evolutionary grounds and by associating substance dualists, including their employment of "introspection," with the Catholic Church, who oppressed Galileo due to fear of scientific progress.[35] The latter approach was employed by the late Lynne Rudder Baker's (and those who follow her) acceptance of a constitutional view of the human person.[36] According to Baker, when a human animal has the property of being-a-first-person-point-of-view, the human animal constitutes a person. One of us has criticized her views elsewhere, but that is not our current concern.[37] If SD is true, the "first-person point of view" is not a property we have; rather, in a certain sense, it is something we are. The "first-person point of view" is identical to a substantial soul having conscious states of a personal kind—especially awareness states—directed toward intentional objects from the indexical "location" of that soul. For example, a personal soul exemplifying the mental property being-an awareness-of-a-red-surface just is for that person to have a first-person point of view. No additional Bakerian property is needed.

---

[35] Churchland, P. (2013). *Matter and consciousness* (3rd ed., pp. 21–28, 120–122). Cambridge, MA: MIT Press.

[36] Baker, L. R. (2000). *Persons, and bodies: A constitution view*. Cambridge: Cambridge University Press, especially Part I.

[37] Moreland, J. P. (2009). *The recalcitrant imago dei* (pp. 131–137). London: SCM Press.

Suppose this is correct. An important methodological approach follows quite naturally. In doing philosophy of mind, the "first-person approach" along with direct knowledge by acquaintance that constitutes it should be the fundamental way we gain insight into the nature of consciousness and the self and into the role this approach plays in various epistemological debates. In our view, the best way to employ this approach is for substance dualists to embrace and do research in light of analytic phenomenology and a consciousness-first methodology, to which we now turn.

11.2.5 Consciousness-First and Analytic Phenomenology

Rather than treating consciousness as something to be explained, as nearly every project in consciousness studies, the SD can start with consciousness itself to explain or assay various things, including what it is to be a subject of consciousness. More specifically, one can employ a specific feature of consciousness itself (e.g., phenomenal unity) to determine the ontology of consciousness subjects, as done in Chapter 6.

A consciousness-first approach acknowledges an important fundamental methodological distinction. Whereas the naturalist privileges third-person data regarding consciousness, the non-naturalist, often a dualist of some variety, gives epistemic primacy to first-person data. The naturalist and non-naturalist start from differing philosophical commitments. As Harold Langsam explains,

> According to the dualist, the physicalist denies the existence of the very essence of the mind. And the dualist gets incensed about this denial because she finds this essential feature of the mind to be so *interesting*, in that it is inextricably intertwined in all kinds of surprising and fascinating ways with almost everything that makes human life interesting and worthwhile.[38]

Langsam urges the non-naturalist to invest her efforts in developing a positive project showing how the non-naturalist conception of the mind is indispensable. The dualist should ignore the naturalist's framing of the debate, which places the non-naturalist on the defensive.

A consciousness-first approach takes an ontologically robust view of consciousness as a starting point from which to explain the nature and possibility of other things, such as intentionality, mental causation, rationality, normativity, and the meaning of our lived experience. Adam Pautz, for example, argues that reductive accounts of consciousness cannot accommodate the various ways in which consciousness is significant (e.g., dissimilarity-grounding, reason-grounding, and thought-grounding), although non-reductive accounts of consciousness can.[39] Further, as noted in Chapter 7, we see no reason to limit phenomenal consciousness to sensations. We think SDists should reject the strategy that does limit phenomenal consciousness this way and functionalizes propositional attitudes. In our view, SDists should do more work defending and exploring the implications that follow from treating sensations, thoughts, beliefs, desires, and exercises of active power as states of phenomenal consciousness.[40] The SDist should expect similar discoveries that are unique to SD.

---

[38] Langsam, H. (2001). Strategy for dualists. *Metaphilosophy*, *32*(4), 396.

[39] Pautz, A. (2017). The significance argument for the irreducibility of consciousness. *Philosophical Perspectives*, *31*(1), 349–407.

Extending Langsam's proposal beyond property dualism, we urge that substance dualists focus on projects of explaining how and why the mind or self—an immaterial, simple substance—is interwoven with central features of reality and how SD is exceptionally well suited for such explanations.

Central to a consciousness-first philosophy of mind is a commitment to a commonsense epistemology, which includes the foundational role knowledge by acquaintance plays in philosophical analysis.[41] It is our direct acquaintance with our own conscious life that provides strong but defeasible evidence for certain commonsense beliefs. Such beliefs plausibly include, for example, that my consciousness is unified, that I am not identical to my stream of consciousness, that my mental states causally interact with the physical states of my body, and that my thoughts can accurately represent reality.[42]

Now Langsam rejects our proposal, arguing that a consciousness-first approach is legitimate only for the property dualist. The consciousness-first program is motivated by the dualist intuition, which Langsam clarifies as an intuition about the nature of a certain property, namely consciousness. It is not an intuition about the nature of the substance that can instantiate this property. The dualist intuition directly supports property dualism, not substance dualism.[43]

According to Langsam, we are directly acquainted with the nature of consciousness but not with the intrinsic nature of the self. This argument is exceptionally weak and empirically and phenomenologically undermined (see Chapter 5). The substance dualist has every right to take a consciousness-first methodology.

Accordingly, an SD research program can assume robust realism about consciousness. One can reject the presumption of naturalism and the burden of defending non-naturalism. SD can take as a starting point the commonsense belief that consciousness is real and that one can and does know certain salient features of consciousness by being directly acquainted with one's consciousness. SD proceeds from the evidence one has and considers consciousness-first starting points as defeasible in the face of more substantial evidence to the contrary if such evidence arises.

---

[40] Pitt, D. (2004). The phenomenology of cognition or what is it like to think that P? *Philosophy and Phenomenological Research*, *LXIX*, 1–36.

[41] This follows the consciousness-first project, as stated by Adam Pautz. See, for example, Langsam, H. (2002). Consciousness, experience, and justification. *Canadian Journal of Philosophy*, *32*(1), 1–28; and Pautz, A. (2013). Does phenomenology ground mental content? In U. Kriegel (Ed.), *Phenomenal intentionality: New essays* (pp. 194–234). Oxford: Oxford University Press.

[42] For recent work on the acquaintance approach to consciousness, studies see, e.g., Smithies, D. (2019). *The epistemic significance of consciousness*. Oxford: Oxford University Press; Gertler, B. (2012). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 93–127). Oxford: Oxford University.

[43] Langsam, Strategy for dualists, p. 379.

As a methodological implication, like Langsam, we suggest that the substance dualist gives primacy to a kind of analytic phenomenology primarily inspired by those such as Franz Brentano, Edmund Husserl, Bernard Bolzano, Gustav Bergman, Reinhardt Grossmann, and Roderick Chisholm.

> *Analytic Phenomenology*: Philosophical inquiry begins from the first-person perspective by considering commonsense beliefs and carefully considered intuitions concerning the nature of phenomenal consciousness, followed by relevant empirical data and appeal to theoretical virtue.[44]

Moreover, the substance dualist should develop research projects that direct our attention to detailed phenomenological descriptions of phenomenal consciousness and, perhaps more importantly, sustained phenomenological descriptions of self-awareness that draw attention to the soul, the immaterial, substantial self, as the subject of consciousness. These projects aim to help one focus one's intentionality in ways that will reveal in direct acquaintance what SD claims is true.

11.3 DEVELOPMENTAL SD RESEARCH PROJECTS

11.3.1 Charting the Logical Space of SD Embodiment

11.3.1.1 Two-Dozen or so SD Embodiment Distinctions

A further way to carve up the SD logical space is to consider views of embodiment that are compatible with SD. We begin with two standard SD views:

> *Compound Substance Dualism*: I am identical to a composite of two parts, body and soul.
>
> *Simple Substance Dualism*: I am identical to a mereologically simple soul, which is now embodied.

Consider the varieties of *Compound Dualism*:

(a) I am identical to a composite of two essential parts (i.e., constituents), a substantial soul and a body.
(b) I am identical to a composite of two essential parts, a substantial soul and a substantial body.
(c) I am identical to a composite of two parts, a soul, and a body, with the former contingent and the latter essential.
(d) I am identical to a composite of two parts, a soul, and a body, with the former essential and the latter contingent.

---

[44] This statement is a highly revised version of Phil Goff's characterization of analytic phenomenology. See Goff, P. (2017). *Consciousness and fundamental reality* (p. 271). Oxford: Oxford University Press.

Similarly, consider variations of *Simple Dualism*:

(e) I am identical to a substantial soul and am essentially embodied.
(f) I am identical to a substantial soul and am contingently embodied.

These qualifications bring to our attention the fact that SD is compatible with various views of embodiment:

Contingent Embodiment: That I am embodied is not essential to my identity.
Necessary Embodiment: That I am embodied is essential to my identity.
Initial Embodiment: My being embodied is essential to my coming into existence.
Normative Embodiment: That I am embodied is not essential to my identity but is essential to my flourishing.
Persisting Embodiment: That I am embodied is not essential to my identity but is nomologically essential to my persistence.
Teleological Embodiment: The powers of my soul teleologically develop my embodiment.

None of these theses are obviously incompatible with standard arguments for SD. They are live options to further develop, perhaps in relation to other embodiment views.

The logical space also permits various modal commitments regarding disembodied existence.[45]

(g) The soul can exist without the body existing.
(h) The soul can exist without being causally related to the body.
(i) It is (logically) impossible that the soul can exist without the body, but the soul can exist without being causally related to the body.
(j) It is (logically) possible that the soul can exist without being causally related to the body but cannot exist without the body existing.

Moreover, there are various ways something could be my body.

(k) The soul can have multiple bodies (synchronically or diachronically).
(l) I can exist as a composite of one soul and multiple bodies (synchronically or diachronically).
(m) What makes *x* my body is a causal relation between *x* and my soul.
(n) What makes *x* my body is a teleological relation between *x* and my soul.
(o) What makes *x* my body is the soul's informing and being holenmerically present throughout my body.

Many of these views can specify various neurological theses. Here are some possibilities:

Neurological Dependence: Mental activity (partially or completely of the soul) is neurologically dependent.

---

[45] The following distinctions are from Meixner, U. (2017). Substance dualism. In Y. Nagasawa & B. Matheson (Eds.), *The Palgrave handbook of the afterlife* (p. 279). London: Palgrave Macmillan.

> *Neurological Independence*: At least some and only some mental activities (partially or completely of the soul) are neurologically dependent.
>
> *Neurological Interdependence*: At least some mental activity is neurologically dependent and soul dependent.

11.3.1.2 SD Embodiment and the Unity of the Person

Contemporary SD work is primarily devoted to sustained criticisms of physicalism and materialism and defending SD. Comparatively, SD work on embodiment is underdeveloped. In a recent article on the revival of SD, Howard Robinson discerns,

> There is, however, something that those of us who believe in mental substance need to say more about, namely the contribution of the immaterial self to the personhood and psychology of human beings ... We need more on what one might call the philosophical psychology of substance dualism, and of how embodiment—the real issue of "the mind–body problem"—might work.[46]

Each view in the previous sections can be developed in conjunction with a specific SD theory. Emergent SD clearly expressed this. Given its neo-Aristotelian commitments, staunch Hylomorphic SD also entails certain robust embodiment and neurological dependence theses.[47]

A longstanding objection is that SD undermines the unity of the human person. This was pressed against Descartes by Arnauld, Gassendi, Princess Elizabeth, and Regius.[48] Descartes insisted that the soul does not merely reside in the body "as a pilot resides in a ship," but instead forms a natural unity with it.[49] He explains

> that the soul is really joined to the whole body, and that we cannot properly say that it exists in one part of the body to the exclusion of the others. For the body is a unity which is in a sense indivisible because of the arrangement of its organs, these being so related to one another that the removal of any one of them renders the whole body defective. And the soul is of such a nature that it has no relation to extension, or to the dimensions or other properties of the matter of which the body is composed: it is related solely to the whole assemblage of the body's organs.[50]

---

[46] Robinson, H. (2021). The revival of substance dualism. *Roczniki Filozoficzne*, *LXIX*(1), 43. Robinson suggests that this project might point toward a more Thomistic picture of the relation of body and soul.

[47] Although a standard Thomistic dualist, the recent work of Daniel De Haan is applicable. See De Haan, D. (2017). Hylomorphism, new mechanisms, and explanations in biology, neuroscience, and psychology. In R. C. Koons, N. J. Teh, & W. M. R. Simpson (Eds.), *Neo-Aristotelian perspectives on contemporary science* (pp. 293–326). New York, NY: Routledge; and De Haan, D., & Meadows, G. A. (2013). Aristotle and the philosophical foundations of neuroscience. *Proceedings of the American Catholic Philosophical Association*, *87*, 213–230.

[48] See, e.g., Arnauld, Fourth set of objections, *CSM* II, 143; Gassendi, Fifth set of objections, *CSM* II 233–8; Descartes, Letter to Regius, December 1641, *CSM* I 200; Descartes, Letter to Princess Elizabeth, May 1643, *CSM* III, 217–20. For a helpful discussion of these objections, see Fowler, D. F. (1999). *Descartes on the human soul* (pp. 340–409). Netherlands: Kluwer Academic Publishers.

[49] René Descartes, *Meditations on First Philosophy*, sixth meditation, 81; also *Discourse on method* 1637, *CSM* I, part 5: 59, 141.

[50] Descartes, The passions of the soul, *CSM* I, 339.

"One can read passages like this," says Susan James, "as attempts to establish that the Cartesian account of the unity of the person is as strong, or even stronger, than its Aristotelian rival."[51]

Our point isn't that Descartes successfully replies to the objection, but rather that substance dualists are well aware of these problems and have attended to them in great detail. But, unfortunately, contemporary literature entirely overlooks these accounts of embodiment

11.3.1.3 The Bodily Soul View

One of us, along with C. Stephen Evans, has defended the *bodily soul view*, according to which each human person is a *bodily soul*, an immaterial simple substance that requires embodiment, not for one's metaphysical identity or persistence, but for one's proper function and flourishing.[52] Unlike some forms of mind–body dualism, my body is not a separable part of me to which I am merely causally linked.

One might ask why, if self and body are so intimately related, we should not simply identify a person with his or her body. The answer is that a person as a self must be distinguished from his or her body. Identity is a necessary relation. If I am identical to my body, then it is necessary that what is true of my body is also true of me and vice versa. However, because a person has some characteristics *qua* self that the person does not have *qua* body, it is not logically possible to identify a person with his or her body.

Embodiment is the mode in which I manifest my presence in the world and exercise my agency and relationality. "To live as a person," says Husserl, "is to posit oneself as a person, to find oneself in, and to bring oneself into, conscious relations with a 'surrounding world.'"[53] As such, the human body plays a dual role. The self is a *bodily self*; thus, my body is not simply another object in the world. It is rather the form in which I exercise my agency and the primary means of how I show up in the world. If I move from point *a* to point *b*, I do so by moving my body. However, the body is also experienced as an object in the world. It exhibits an indifference to the world. If my legs are trapped under a car, I cannot move from point *a* to point *b*. If a tumor invades my brain's primary motor cortex, the neural signals needed for bodily movement will not be generated and I will be unable to move.

I thus find myself thinking of my body in two distinct ways: both as the locus of my agency; the mode in which I exist as a conscious self, and as an object in the world; a psycho-physical unity often under the control of my will. When we think of the body in this second way, we naturally think of it as something distinct from our self; we think of the body as if it were another object in the world, an entity whose characteristics I must consider when I act.

---

[51] James, S. (2000). The emergence of the Cartesian mind. In T. Crane & S. Paterson (Eds.), *History of the mind-body problem* (p. 126). London: Routledge. Cf., Koivuniemi, M., & Curley, E. (2015). Descartes on the mind–body union: A different kind of dualism. In D. Garber & D. Rutherford (Eds.), *Oxford studies in early modern philosophy* (Vol. VII, pp. 83–112). Oxford: Oxford University Press.

[52] See, Rickabaugh, B., & Stephen Evans, C. (2018). Neuroscience, spiritual formation, and bodily souls: A critique of christian physicalism. In K. Loftin & J. R. Farris (Eds.), *Christian Physicalism?* (pp. 231–56). Lanham, MD: Lexington Books; and Evans, C. S., & Rickabaugh, B. (2015). What does it mean to be a bodily soul? *Philosophia Christi*, *17*(2), 315–330.

[53] Husserl, E. (1989). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy—Second book: Studies in the phenomenology of constitution* (R. Rojcewicz & A. Schuwer, Trans., p. 193). Dordrecht: Kluwer Academic Publishers. First published 1952.

When I think of my body as the mode in which I exist as a self, it is not a mere object, but myself. When I think of my body in this second, objectified manner, however, it is natural to think of it not as myself, but as something that the self must take into account in its agency, a part of the physical world. When I think of the body in this objectified way, it is natural to think of it as something distinct from the self. Hence, the language of body and self or body and soul as two distinct entities is not only appropriate because of their distinct properties; it is also appropriate insofar as we conceive of the body in this objectified manner. I am, at the same time, wholly soul and yet fully embodied. As Ludwig Wittgenstein observes, "The human body is the best picture of the human soul."[54]

11.3.1.4 Staunch Hylomorphic SD

While the association of hylomorphism with SD might seem bizarre, it is not without precedent. Marmodoro and Cartwright observe,

> Some late antique thinkers depart very explicitly from Aristotle's treatment of the soul–body relation as part and parcel of his general hylomorphic account of reality; rather, they take an anti-naturalist stance and embrace some version of substance dualism whilst also espousing a version of hylomorphism.[55]

The following sketch has been and is especially influenced by the late medieval Aristotelians,[56] Bernard Bolzano,[57] E. J. Lowe,[58] A. G. A. Balz,[59] among others.[60]

On the proposed view, the self or soul is ontologically fundamental. Facts about the soul ground facts about the body. It is here that certain hylomorphic theses are helpful.

---

[54] Wittgenstein, L. (1953). *Philosophical investigations* (3rd ed., G. E. Anscombe, Trans., p. 178e). New York, NY: The Macmillan Company.

[55] Marmodoro, A., & Cartwright, S. (2018). Introduction. In A. Marmodoro & S. Cartwright (Eds.), *A history of mind and body in late antiquity* (p. 4). Cambridge: Cambridge University Press.

[56] See, e.g., Pasnau, R. (2011). *Metaphysical themes: 1274–1671*. Oxford: Clarendon, for example, 549, 558, 560–565; and Des Chene, D. (2000). *Life's form: Late Aristotelian conceptions of the soul* (pp. 67–113; 191–199). Ithaca, NY: Cornell University Press.

[57] Bolzano, B. (1838). *Athanasia; oder Gründe für die Unsterblichkeit der Seele* (pp. 55–56, 101, 283–284). Sulzbach: J. G. v. Seidleschen Buchhandlung.

[58] Lowe, E. J. (1996). *Subjects of experience* (pp. 14–51). New York, NY: Cambridge University Press; and Non-Cartesian mind-body dualism. (2012). In B. P. Göcke (Ed.), *After physicalism* (pp. 48–71). Notre Dame, IN: University of Notre Dame Press.

[59] Balz, A. G. A. (1951). *Cartesian studies* (pp. 279–323). New York, NY: Columbia University Press.

[60] See, e.g., Voss, S. (1992). Understanding eternal life. *Faith and Philosophy*, *9*(1), 3–22; Barnes, G. (2004). Is dualism religiously and morally pernicious? *American Catholic Philosophical Quarterly*, *78*(1), 99–106; Barnes, G. (2002). Should property-dualists be substance-hylomorphists? *Proceedings of the American Catholic Philosophical Association*, *75*, 285–299; Schärtl, T. (2012). Beyond dualism? The track-switch model of resurrection. In B. P. Göcke (Ed.), *After physicalism* (pp. 335–368). Notre Dame, IN: University of Notre Dame Press.

Like Augustine,[61] Aquinas,[62] and Suarez,[63] the proposed view embraces Aristotle's thesis that the soul is "the cause and source of the living body"[64] As Edmund Husserl says, "the soul … be-souls the Body."[65] Alternatively, on Bolzano's powers ontology, the soul "rules over" the body by the comparatively superior "strength" of its powers, which interact with the powers of the body's material simples.[66] Consequently, the human body cannot exist without the soul, although the soul can exist without the body. Here is a standard illustration.

> *Salted Water*: Suppose I have a glass of water and a spoonful of salt. When I put the salt into the water, the salt dissolves, and I now have a glass of saltwater. If I were able to remove the salt from the water, I would have plain water in the glass and salt on the table.

The salt represents a human soul, the salted water represents a human body, and the water left over once the salt is removed is a human corpse. The salted water cannot exist without the salt, although the salt can exist without the water. Likewise, a human body cannot exist without a human soul, although the soul can exist without the body.

Furthermore, facts about the soul's essence determine, among other things, the soul's natural kind.[67] The soul contains, as a primitive unity, various powers for developing the body. These powers of the soul ground both the internal (e.g., immanent) efficient cause of and teleological guide for the internal structure and development of the body.[68] Certain powers of the soul act as both the first efficient cause of the body's development and the final cause of the body's functions and structure. The soul, not Humean laws, gives, for example, the electrons in my body their identity and characteristic patterns of behavior.

Although largely unrecognized in contemporary literature, substance dualists have thought of the soul as that which forms and sustains the human body as a unified, ensouled biological entity.[69] Thomas Reid is one such unacknowledged example.[70] We can also understand Leibniz as a kind of dualist who embraced similar hylomorphic notions. In reply to Arnauld, Leibniz explains that "Our body is the matter and our soul the form of our substance."[71] Leibniz, like Aristotle, thinks this is true of all living bodies.[72] That views such as Reid's and Leibniz's are largely unexplored in contemporary literature means that treatments of SD are incomplete and ripe for development and application.

---

[61] Augustine, *On freedom of the will*, II, XVI, 41; *The Trinity*, IV, I, 3; *Confessions*, II, 6.

[62] See, e.g., Thomas Aquinas, *Summa theologiæ*, I, Q75; and Aquinas, T. (1949). *Disputed questions on spiritual creatures* (M. C. Fitzpatrick & J. J. Wellmuth, Trans.). Milwaukee: Marquette University Press, IV.ad 9.

[63] Suárez, F. (2000/1597). *On the formal cause of substance: Metaphysical disputations XV* (J. Kronen & J. Reedy, Trans., pp. 140–142). Milwaukee, WI: Marquette University Press.

[64] Aristotle, *De anima*, 415b.9. See also, *De anima* 412a22, 412a27-28.

[65] Husserl, E. (1993). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy, second book: Studies in the phenomenological constitution* (R. Rojcewicz & A. Schuwer, Trans., p. 185). Dordrecht: Kluwer Academic Publishers.

[66] See Lapointe, S. (2018). Bolzano's philosophy of mind and action. In S. Lapointe (Ed.), *Philosophy of mind in the nineteenth century: The history of philosophy of mind* (Vol. 5, pp. 49–53). New York, NY: Routledge.

[67] A natural kind is understood here as akin to a set without extensional identity conditions. It is a collection (or potential collection) of all and only members whose essence is the same and which grounds their natural-kind membership.

[68] See Pasnau, pp. 549, 558, 560–565.

In doing further research in light of a staunch hylomorphic SD, one of the first items of importance is to abandon the ubiquitous and rather tiresome employment of appeals to emergent properties and replace them with the appearance of new substances due to substantial changes. The former is a part of micro-physical priority, and the latter is based on the ontological independence of substances. Among other things, if a certain change produces new kinds of properties, then a SD could look to see if there are signs of a new kind of unity (the whole ontologically prior to its inseparable parts; the inseparable parts are internally related to each other or the whole of which they are parts, etc.) and of irreducible teleology where that is relevant. Given the rich ontological framework within SD, we have no doubt that others will discover exciting new implications for research that follow most naturally from that ontology and the SD being envisioned.

11.3.2 The Great Importance of Embodiment for SD

Taliaferro defends what he considers a completely integrated SD understanding of embodiment according to which the embodied human person is functioning in a healthy way as a unity, what he calls *Integrative Dualism*.[73] Taliaferro depicts the embodiment of a human person as involving the coordination and function of at least six interrelated nonmoral goods or virtues.[74]

> *Sensory*: A materially embodied human person's ability to experience an array of embodied sensations: visual, auditory, olfactory, gustatory, tactile, vestibular, and proprioception.

---

[69] In fact, some Descartes scholars argue that a careful reading of the overall work of Descartes reveals that his view is much closer to late medieval hylomorphists than commonly acknowledged. Arguably, Descartes' view is more comparable to scholastic-Aristotelian theories of soul-body union than is popularly portrayed. See, e.g., Skirry, J. (2005). *Descartes and the metaphysics of human nature*. London and New York: Thoemmes-Continuum Press; Rozemond, M. (2003). Descartes, mind-body union, and holenmerism. *Philosophical Topics*, *31*(1/2): 343–367; Rozemond, M. (1998). *Descartes's dualism*. Cambridge: Harvard University Press, and Hoffman, P. (1986). The unity of Descartes' man. *The Philosophical Review*, *95*, 339–369.

[70] See, Rickabaugh, B. (work in preparation). Discovering Thomas Reid's hylomorphic substance dualism.

[71] Leibniz, G. W. (1902). Leibniz to Arnauld, October 6, 1687. In G. W. Leibniz, *Discourse on metaphysics, correspondence with Arnauld, and monadology* (G. Montgomery, Trans.). La Salle, IL: Open Court, GP 2: 119.

[72] All living bodies are such because "a soul, or form analogous to the soul, or a first entelechy, or a primitive force for acting, which is a law inscribed in it, impressed upon it by divine decree." Leibniz, ibid., 4: 512.

[73] We take it that the kind of unity Taliaferro is after is compatible with, but not committed to, the ontology of holistic unity we explored in Chapter 3.

[74] Taliaferro, C. (2001). The virtues of embodiment. *Philosophy*, *76*, 111–125. Taliaferro understands virtue as an excellence in respect to a thing's nature, operation, worth, or efficacy of any kind.

> *Epistemic*: A materially embodied human person's power to be aware of one's body in an epistemically justified or warranted way.
>
> *Structural*: A materially embodied human person's ability to exercise mental powers in a sustained fashion and to reflect, think, feel, and act in ways that are consistent and interrelated.
>
> *Constitutional*: A materially embodied human person's powers that sustain psychological life, the powers that enable one to be a Lockean person as an embodied, human being.
>
> *Affective*: A materially embodied human person's self-respect (at least a minimal level of) in which one has affectively incorporated one's body as one's own.
>
> *Agency*: A materially embodied human person's being able to affect one's body directly, by one's intending to do so.

This understanding of embodiment, argues Taliaferro, provides SD with a strong link between soul and value theory, an account on which a substantial soul requires a body to flourish.

In a similar approach, Richard Swinburne argues that embodiment is required on SD if we have choices between good and evil, which is a great good we should have.[75] Swinburne offers four conditions on the kind of embodiment we have right now involves:

> *Control*: A (normally moveable) public region of basic control,
>
> *Action/Perception*: A (normally moveable) public region of basic action and perception—regions through which one is tied down to acting and perceiving,
>
> *Intention*: A public "machine room" necessary for translating one's intentions into one's actions and perceptions, which
>
> *Desire*: Is also the source of our desires.

To the kind of morally relevant choice, Swinburne argues, requires all four embodiment conditions for being like us as we currently are. I need some power over and knowledge of a region that is different from those of others. My regions of basic control and perception must be limited if I am to have the intentional choice of whether or not to extend them. My regions of control and perceptions must belong to the same public world as others if I (and we) am to have responsibility for communication (of our beliefs and desires) and cooperation with others. And among other things, my "machine room" must be accessible to and predictable by others for them to have a choice about causing me pleasure or pain and, conversely, for me in relation to others. Each is, among others, necessary for the great good of having the power to choose for good or evil, which also requires our having a body on an SD view.

Other related research should be developed on issues such as the value and significance of gender (is the soul gendered or just the body? If so, does the sex/gender of the soul need to match the sex/gender of the body? What implications, if any, do the answers to these questions have for transgender questions?), the value and significance of social kinds (is my soul or just my body, or neither, racial?), and finally, issues concerning disability (is my soul disabled or just my body or both or neither or it depends, etc.?).[76]

---

[75] Swinburne, R. (2003). What is so good about having a body? In T. W. Bartel (Ed.), *Comparative theology: Essays for Keith Ward* (pp. 134–142). London: SPCK Publishing.

[76] We wish to thank Robert Garcia and Alexander Pruss for these suggestions.

Although much more work is needed, substance dualists have provided accounts of embodiment and its great importance. In addition to developing these accounts further and proffering alternatives, substance dualists must draw attention to these accounts and the fact that they remain unchallenged.

11.3.3 SD and Psychological Development

Some argue that SD faces problems accounting for human psychological development. In the words of Kevin Corcoran,

> If what dualists say about the soul is true — that it's a nonphysical, partless entity and that it is the bearer of psychological properties and states of human consciousness — then, it seems difficult, if not impossible, to account for the gradual development of conscious capacities, as it's hard to see how a partless thing can develop at all. In other words, it would seem that the soul that is you must have had the capacity to engage in the solving of differential equations from its earliest days as the soul of the embryo that became your body. Yet we know that an embryo's neural structure lacks the complexity necessary to underwrite the cognitive sophistication one needs to think about, let alone solve differential equations.[77]

Regardless of how often this mantra is repeated, it is demonstrably wrongheaded and a strawman account of SD.[78] For one thing, Corcoran affirms the mistaken notion—latter asserted by Jaegwon Kim—that a simple (unextended) soul would have no internal structure, especially the structure required to develop gradually the capacities of complex thought.[79] Clearly, Corcoran is employing only separable, and not inseparable parts, in his critique. Accordingly, *sans* (separable) parts, the soul is partless in any sense relevant to development. Moreover, if one reflects on the various properties that constitute consciousness, especially the capacities for such, it becomes evident that the possessor of these capacities and properties do not require a spatially extended subject to be exemplified. For many, if not most, substance dualists, the soul contains a rich internal structure comprised of various exemplified properties and capacities (dispositions or powers), faculties, and a multiplicity of causal interactions and dependency relations between and among them. Corcoran's (and Kim's) ontology is too undeveloped as evidenced by the fact that such an uninformed critique would be raised as problematic in the first place.

There is another problem with this Corcoran-style objection. It fails to understand the hierarchically ordered arrangement of the soul's capacities and the soul's dependence on the organs and systems of the brain and body and for their exercise and development. This hierarchical arrangement has first-order, second-order, and ultimate-order capacities that constitute the soul's essence. Among other things, development involves the ultimate-order capacities unfolding as the various ordered capacities under it are actualized. If for some reason, there is a defective blockage at, say, the level of a first or second-order capacity, the soul retains the relevant ultimate capacity by nature but cannot exercise it due to blockage at a lower level. Some of this blockage is merely the developmental stage at which the brain and other relevant bodily constituents exist. Another way to see this is to note that many substance dualists take the exercise of many, if not most, of the soul's powers to be dependent upon the relevant features of the body required (while embodied) to be employed. There is very little work on SD views of psychological development.[80] This is a great need worthy of sustained attention.

---

[77] Corcoran, K. (2018). Why should a Christian embrace materialism? In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (p. 288). Oxford: Wiley-Blackwell.

[78] See, e.g., O'Connor, T., & Wong, H. Y. (2005). The metaphysics of emergence. *Noûs*, *39*(4), 660. This argument is attributed to Armstrong, D. (1968). *A materialist theory of mind* (pp. 30–31). London: Routledge and Kegan Paul.

[79] See Kim, J. (2005). *Physicalism, or something near enough* (pp. 88–90). Princeton: Princeton University Press, especially 90.

Turning from a critique to a positive way forward, we begin by noting that in 2003, Jeffrey M. Schwartz and Sharon Begley published a revolutionary new treatment for various mental disorders, especially OCD, that resulted from a distinct SD research program.[81] The authors claim that prior to their research, a behavior-modification technique based on a physicalist ontology of the human person was widely—often traumatizing and ineffective—used as a treatment strategy. However, by postulating a substantial, immaterial mind with the powers of libertarian freedom, they developed and successfully tested a new therapeutic approach that was highly successful. This is a clear case in which SD assumptions generated empirically testable results and were, in the authors' words, confirmed.

We urge substance dualists to develop more research like this, perhaps by SD philosophers and neuroscientists working together.

11.3.4 SD and the Nature of the Physical

Those involved in debates between dualists of various kinds and physicalists would do well to consider the role that theories of the physical play in their disputes. The assumption is that they share and must share the same theory of matter. But this needn't be the case. Why can't the substance dualist posit a range of powers in the physical that can uniquely be tied to the powers of the soul? Why can't the substance dualist utilize a version of hylomorphism? Doing so would undoubtedly impact issues of mind–body causation, perception, and the unity of subjects. We suggest that these issues be explored and applied to both new views of SD and longstanding debates over SD.

Moreover, we agree with Goetz and Taliaferro. They urge substance dualists to draw greater attention to the fact that we simply do not have a decent understanding of the physical, which poses problems for physicalism.[82] The current opacity of the physical is also an opportunity for alternative SD-friendly theories of the physical. Such opacity also changes the employment of the *via negative* of characterizing the soul (see Chapters 8 and 10). The standard approach is to take our notion of the physical as fairly straightforward and, starting with that notion, to go on to define the soul in purely negative terms, e.g., an immaterial substance. We recommend that SDists discard this approach and replace it with one according to which the nature of the soul is clearer than that of (dispositionally and formally characterized) physical entities. Accordingly, SDists should characterize the soul positively in terms of (1) the formal features of a substance and (2) the nature of properties of phenomenal consciousness (and, for some, the powers of life).

---

[80] Here is one exception: Robinson, H. (1987). A dualist perspective on psychological development. In J. A. Russell (Ed.), *Philosophical perspectives on developmental psychology* (pp. 119–139). Oxford: Basil Blackwell.

[81] Schwartz, J. M., & Begley, S. (2003). *The mind and the brain: Neuroplasticity and the power of mental force*. New York, NY: Harper Perennial.

[82] Goetz and Taliaferro, *A brief history of the soul*, pp. 205–210.

11.4 APPLIED SD RESEARCH PROJECTS

SD has been applied to various issues outside of the philosophy of mind and the ontology of human persons. Here we suggest further areas of research and application. Some projects are mere suggestions, while others are more developed.

11.4.1 SD and New Waves in the Philosophy of Religion

Perhaps the most obvious SD research projects are about issues in the philosophy of religion. What follows is primarily a list of brief suggestions for further inquiry.

SD and knowledge by acquaintance of the divine. As we've shown throughout this book, there has been a strong return to knowledge by acquaintance and a growing literature in second-person and interpersonal knowledge in the philosophy of religion.[83] We think the SDist has something to offer. For example, the kind of knowledge by acquaintance that we defended in Chapter 5 opens up the door to defend more uniformity in experiences of God in daily life. By making an interpretation or conceptualization of the experience derivative and not constitutive of the experience by taking everything as seeing as, one can claim that what people actually experience throughout various situations has a lot in common. Still, when they interpret and report the experience, often, they bring their preexisting concepts to the task, and one gets diversity of interpretative reports even though there was more uniformity. Maybe they all experienced God but interpreted the object of experience in different ways. How might this impact debates about religious disagreement and religious pluralism?

Divine mental causation. How might SD work on the objections from mind–body interaction to help illuminate issues of divine causation, creation, answered prayer, and miracles?[84] Or, if theism is true, might the reality of divine mental causation make the mind–body causation a non-issue? Perhaps there is an argument for atheism lurking in the background?

---

[83] See, e.g., Benton, M. A. (2018). God and interpersonal knowledge. *Res Philosophica*, *95*(3), 421–447; Rickabaugh, B. (2013). Eternal life as knowledge of god: An epistemology of knowledge by acquaintance and spiritual formation. *Journal of Spiritual Formation and Soul Care*, *6*(2), 204–228; and Stump, E. (2010). *Wandering in darkness: Narrative and the problem of suffering*. Oxford: Clarendon Press.

[84] See, e.g., the chapters by Angus Menuge and Gregory E. Ganssle in Ganssle, G. (Ed.). (2022). *Philosophical essays on divine causation*. New York, NY: Routledge; and Lim, D. (2015). *God and mental causation*. New York, NY: Springer.

Intentionality and mindfulness or meditative practices. Mindfulness has as a necessary condition, the having of second-order conscious states that are irreducibly *of* the intrinsic phenomenal content, the content of first-order conscious states.[85]

Consciousness-first natural theology. The consciousness-first project in the philosophy of mind can be applied to natural theology in a way similar to how we have applied it to SD. Might the theist be justified, just as the property dualist or SDist likely is, in taking consciousness (human and divine) as not something to be explained but as something to explain other things? Perhaps the fundamentality of consciousness or, better yet, subjects of consciousness can be applied to arguments for the existence of God.

Accounting for empirical research on NDEs. Many reported near-death experiences (NDEs) studies make it clear that people immediately experience themselves as completely disembodied selves or I's. Yet, although they experience being invisible, they have a spatial location and "point of view" from somewhere in the room. When they cross over to the other side (e.g., via going thru a tunnel), the intermediate state seems physical—trees, rivers, flowers, etc. Further, they seem embodied. This might seem problematic for SD, suggesting that there is no out-of-body experience. SDists should work on what this means. For example, if they are re-embodied, SD still provides a grounding of personal identity for two different bodies. Second, when their bodies are described, they seem fuzzy (sort of spread throughout an extended location yet more like a cloud without distinct, precise boundaries), they often seem to have bodies made of light, etc. Third, maybe the intermediate state should be interpreted with some version of Berkeleyan Idealism.[86]

Divine ideas and universals. Debates over divine ideas and universals have gained some popularity over the past fifteen years. There may be reasons for the SDist to think that universals cannot be concepts in God's mind or God's thoughts. First, concepts are themselves mental/intentional universals. Second, if God's mind exemplifies redness, his mind is spatially extended and colored red. Medieval philosophers tried to avoid this by saying, e.g., redness is exemplified in a diminished way in minds, so it doesn't make the mind red. This seems *ad hoc*. Third, concepts are irreducibly about their objects; they are not identical to them.

Naturalist SD. Against the standard narrative, it must be acknowledged that SD does not strictly entail theism. As explored in Chapter 1, atheists and agnostics such as C. D. Broad, Evan Fales, Michael Huemer, Michael Martin, Graham Oppy, Karl Popper, and Eric Steinhart defend SD.[87] An interesting project would work out the ontological details of a quasi-naturalistic or mere-atheistic SD. An excellent place to start would be with recent arguments for theism from consciousness.[88]

---

[85] See, e.g., Siegel, D. J. (2010). *Mindsight: The new science of personal transformation*. New York, NY: Bantam Books.

[86] Examples of such projects are sparse. See, e.g., Stafford Betty, L. (2004). Mind, paranormal experience, and the inadequacy of materialism. *International Philosophical Quarterly, 44*(3): 373–392; and Potts, M. (2001). Sensory experiences in near death experiences and the Thomistic view of the soul. *International Journal for Philosophy of Religion*, *49*, 85–100.

[87] See, Broad, C. D. (1959). Autobiography. In P. A. Schilpp (Ed.), *The philosophy of C.D. Broad* (pp. 3–68). New York: Tudor; see, 43–44; Fales, E. (2007). Naturalism and physicalism. In M. Martin (Ed.), *The Cambridge companion to atheism* (p. 127). New York, NY: Cambridge University Press; Fernandes, P., & Martin, M. (2000). *Theism vs. atheism: The internet debate* (p. 19). IBD Press; Huemer, M. (2018). *Paradox lost: Logical solutions to ten puzzles of philosophy* (p. 235). Switzerland: Palgrave Macmillan; Oppy, G. (2013). *The best argument against God* (p. 55). New York: Palgrave Macmillan; and Steinhart, E. (2019). Naturalism. In G. Oppy (Ed.), *A companion to atheism and philosophy* (p. 159). New York, NY: Wiley-Blackwell; Zerin, E. (1998). Karl Popper on God: The lost interview. *Skeptic*, *6*(2), 47–48.

Consciousness and anthropic fine-tuning. Consider the anthropic fine-tuning debates. With the growing interest in panpsychism, a non-theistic personal explanation, a self-designing universe, may seem more plausible than ever. Thomas Nagel suggests, "The universe has become not only conscious and aware of itself but capable in some respects of choosing its path into the future."[89] More explicitly, Philip Goff defends a self-designing hypothesis, according to which anthropic fine-tuning is best explained by a universe with a basic form of consciousness, agency, and rationality.[90] Like theists, Goff recognizes that a mindless origin of anthropic fine-tuning is improbable. Goff's proposal advances fine-tuning debates beyond competing personal and sub-personal explanations to competing personal explanations of fine-tuning: a finite, contingent mind or an infinite, necessary mind—a god-like universe or a God of the universe. Of course, SDists will likely reject Goff's thesis.[91]

Our point is that the philosophy of religion has yet to take advantage of the resources of the philosophy of mind in any substantial way. Doing so ought to be fruitful, and SDists are uniquely positioned to produce new SD research projects in this area.

11.4.2 SD and Epistemology

11.4.2.1 Internalism and Externalism

Plausibly, externalism in epistemology is linked with naturalism and materialism in the philosophy of mind, while internalism about justification is related to SD. As Richard Fumerton says, "There is no epistemological Switzerland in the war between dualists and physicalists."[92] This relationship has not been widely acknowledged, nor has its implications.

In both its reductive and non-reductive forms, physicalism makes attractive epistemological externalism about epistemic justification (and, ultimately, knowledge). Much of the externalist program attempts to explain how sub-personal mechanisms or relations, like non-normative causation and reliability, can account for non-inferential justification (or knowledge). However, internalists typically understand non-inferential justification in terms of person-level mental states or processes (e.g., Russell's direct acquaintance, Chisholm's "more reasonable than," Bonjour's introspective infallibilism or the "seemings" of Heumer, and Swinburne, and others). The robust role internalists attribute to mental states or processes (accessible and or phenomenal) in generating epistemic justification seems to run counter to the restricted causal role ascribed by physicalism. Thus, there is a noteworthy, and as of yet unaddressed, *prima facie* conflict between physicalism in philosophy of mind and internalism in epistemology. Likewise, we anticipate other interesting epistemic implications for substance dualists to explore.

---

[88] See, e.g., Adams, R. M. (1987). *The virtues of faith* (Oxford: Oxford University Press,); Moreland, J. P. (2008). *Consciousness and the existence of God*. New York, NY: Routledge; O'Connor, T., & Kimble, K. (2011). The argument from consciousness revisited. In J. Kvanvig (Ed.), *Oxford studies in the philosophy of religion* (Vol. 3, pp. 110–141). Oxford: Oxford University Press; Page, B. (2020). Arguing to theism from consciousness. *Faith and Philosophy*, *37*(3), 336–362; Swinburne, R. (2018). The argument from colors and flavors: The argument from consciousness. In Walls & Dougherty (Eds.) (pp. 293–303); Vandergriff, K. (2018). Naturalism, theism, and multiply realizable mental states. *Religious Studies*, *54*, 91–105. Cf., Oppy, G. (2011). Review of J.P. Moreland's the argument from consciousness: A theistic argument. *European Journal for Philosophy of Religion*, *3*, 193–212.

[89] Nagel, T. (2012). *Mind and cosmos* (p. 124). Oxford: Oxford University Press.

[90] Goff, P. (2019). Did the universe design itself? *International Journal for Philosophy of Religion*, *85*(1), 99–122.

[91] Rickabaugh, B. (work in preparation). The prior structure of consciousness: Franz Brentano's teleological argument contra cosmopsychism.

[92] Fumerton, R. (2013). *Knowledge, thought, and the case for dualism* (p. 91). Cambridge: Cambridge University Press.

One of us has argued for a significant link between externalism and physicalism on the one hand and internalism and dualism on the other by examining the structural similarities between supervenience arguments against physicalism and arguments in defense of epistemic internalism. Consequently, to the extent that this conflict resists resolution, then to that extent those who find themselves attracted to epistemic internalism have a reason to prefer metaphysical dualism (either property dualism or mind–body dualism). At the same time, those attracted to epistemic externalism have a reason to embrace physicalism. This reveals that those who embrace epistemic externalism and SD, for example, Alvin Plantinga, face a severe dilemma, as it seems that one can embrace epistemic externalism or metaphysical dualism, but not both.[93]

Here is one clarification of this last claim. Given the importance of direct knowledge by acquaintance in arguing for SD, the stage is set to develop non-doxastic internalist views of justification and knowledge. To be sure, some versions of externalism regarding knowledge (e.g., Plantinga's) utilize various seemings, including seeming to be directly aware of sense-perceptible objects, in their account. But our view is that such seemings and direct awarenesses are epistemically relevant *per accidents* and not *per se*. Thus, an experiential awareness of a tree does not in itself confer positive epistemic status on the claim that "I see a green object with such and such shape." Instead, it is because such an awareness is a proper constituent of a reliable process or proper causal chain that it confers positive epistemic status to the proposition. Its being a proper constituent in the appropriate process, not its intrinsic features that are possibly directly present to the subject's internal awareness, is what endows the awareness with its epistemic importance.

11.4.2.2 Social Epistemology

If SD is true, then we have reason to think that a group of minds, unless from it emerges a new mind, cannot have irreducible epistemic properties. For example, suppose the phenomenal unity argument for SD in Chapter 6 is correct. In that case, no social group—no aggregate of individual subjects—can be a subject of any irreducible epistemic properties—mental states of a subject. A practical implication in the *philosophy of law* is that corporations cannot be held morally responsible, but only the corporation's members, e.g., administrative leaders or the board of trustees, can be held morally accountable. We suspect there is much more to say about this.

---

[93] Rickabaugh, B. (work in preparation). Epistemology, mind, and hard choices: Externalism or dualism, but not both.

11.4.3 SD and Content Externalism

In our view, substance dualists will and should reject content externalism or wide mental content. Interestingly, Colin McGinn provides the reason why. He observes,

> I believe it is correct to say that [content] externalism, once it is forcefully and vividly stated, strikes many people as a strange and bizarre theory—a theory that simply couldn't be true. It seems to these people to fly in the face of some deeply rooted conception or picture of the mind and its states. The mind, it is felt, just could not be such as to satisfy externalism; it is simply not the kind of thing that externalism represents it as being. My suspicion, listening to such people, is that some quite fundamental assumption about the mind is being threatened by externalism …[94]

According to McGinn, this fundamental assumption is a ubiquitous Cartesian dualist instinct. Thus, he claims, "I want to suggest, putting it crudely at first, that it is the idea that the mind is a substance, that mental states are states of a mental substance."[95] It is this tendency to depict the mind (or soul) as a substance that conflicts with content externalism.

To unpack this a bit, strictly speaking, supervenience is merely a set of conditions (usually) regarding property covariance. But the conventional use of supervenience includes some sort of dependence relation between the supervenient and the subvenient, e.g., modal necessitation or standard ontological dependency. So let us accept this convention. An important reason for employing supervenience is to secure a physicalist ontology. Suppose strict physicalism fails and one must quantify over emergent, non-physical entities. In that case, as long as those entities can be depicted as arising from and dependent on a physical subvenient base, this is sufficient for many physicalists. This is because all the non-physical facts, properties, and so forth are fixed by the physical facts, etc. However, Weak or more preferably, Strong Property Supervenience fails to fix supervenient relational facts merely by populating the subvenient physical base with particulars and monadic properties. For example, such a base is insufficient to fix an object's weight. To accomplish this task, one must appeal to Global Supervenience.[96]

Now it is precisely this move to Global Supervenience that is needed to fix external mental content. This is because weight, being among the tallest, and external mental content are relational states of affairs and not mere particulars exemplifying monadic properties. We now see why McGinn identifies SD intuitions as resistant and contrary to content externalism. For substance dualists, the self is a spiritual, simple substance; it exemplifies intrinsically monadic mental properties to form particular mental states/events. SD is a substance-attribute ontology regarding the metaphysics of the self and consciousness. Substance dualists will resist the shift from substances to relational structures and, along with this, the reality of external mental content.

A helpful research project would be for substance dualists to criticize and offer alternative solutions to a variety of thought experiments (e.g., water thoughts vs. twater thoughts on Twin Earth, arthritis vs. tharthritis) that are taken to support content externalism.

---

[94] McGinn, C. (1989). *Mental content* (p. 14). New York, NY: Basil Blackwell. Brackets ours.

[95] Ibid., p. 15.

[96] See Jackson, F. (1998). *From metaphysics to ethics* (pp. 6–14). Oxford: Oxford University Press.

11.4.4 SD and Social Psychology

SD has been applied to various issues in psychology,[97] cognitive science,[98] neuroscience,[99] political science,[100] bioethics,[101] and business practices.[102] By taking SD seriously, we will likely progress toward understanding human psychology, sociology, and cultural development. As argued in Chapter 5, belief in SD is ubiquitous across the spectrum of human history and cultures. The standard approach is to dismiss the relevance of SD beliefs or reduce them to sub-cultural and sub-personal mechanisms. But this approach disregards the importance of human phenomenology, which we have argued in Chapter 5, at its core unflinchingly SD.

11.4.5 SD Healthcare and Mental Health

A fundamental issue in medicine is the debate about the best criterion for death.[103] Ontologically, substance dualists will be inclined to characterize death as the irreversible departure of the soul.[104] However, substance dualists divide regarding the fundamental epistemic criterion of death. Cartesians and those who hold that the soul is characterized by the possession of the actualization or dispositions of consciousness will employ a higher-brain or whole-brain criterion according to which there is the irreversible loss of whole or higher-brain functioning.

---

[97] See, e.g., Alanen, L. (2008). Descartes' mind-body composites, psychology, and naturalism. *Inquiry, 51*(5), 464–484; Antonietti, A. (2008). Must psychologists be dualists? In A. Antonietti, A. Corradini, & E. J. Lowe (Eds.), *Psycho-physical dualism today: An interdisciplinary approach* (pp. 37–67). Lanham, MD: Rowman and Littlefield; Boyd, J. H. (1994). *Soul psychology: How to understand your soul in light of the mental health movement*. Cheshire, CT: Soul Research Institute; Hall, T. W. (1998). The soul or substantive self as experiencer, actualizer, and representative in psychoanalytic theory. *Journal of Psychology and Theology*, *26*(1): 55–65; Moreland, J. P. (1998). Restoring the substance of the soul to psychology. *Journal of Psychology and Theology*, *26*(1), 29–43; Rickabaugh, B., & Stephen Evans, C. (2018). Neuroscience, spiritual formation, and bodily souls: A critique of Christian physicalism. In Loftin and Farris (Eds.), *Christian physicalism? Philosophical theological criticisms* (pp. 231–256). Lanham, MD: Lexington Books.

[98] See Chapter 6.

[99] LaRock, E. (2016). Neuroscience and the hard problem of consciousness. In T. M. Crisp, S. Porter, & G. A. Ten Elshof (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology* (pp. 151–180). Grand Rapids, MI: Eerdmans; Manzotti, R., & Moderato, P. (2014). Neuroscience: Dualism in disguise. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 81–97). New York, NY: Routledge; Moreland, J. P. (2012). Christianity, neuroscience, and dualism. In J. B. Stump & A. G. Padgett (Eds.), *The Blackwell companion to science and Christianity*. Oxford: Blackwell Publishing.

[100] See, e.g., Friberg-Fernros, H. (2020). The explanatory power of the soul: Why liberals should seriously consider dualism. *Journal of Religious Ethics*, *48*(1), 101–121.

[101] Moreland, J. P., & Rae, S. B. (2000). *Body and soul: Human nature and the crisis in ethics*. Downers Grove, IL: InterVarsity Press.

[102] Cooper, M. J. (2007). Are we sending mixed messages? How philosophical naturalism erodes ethical instruction. *Journal of Business Ethics*, *75*, 171–180; and Michel, A. (2015). Dualism at work: The social circulation of embodiment theories in use. *Signs and Society: Supplement*, *3*(S1), S41–S69.

[103] See Settimio, & Rid, A. (2012). Controversies in the determination of death: Perspectives from Switzerland. *Swiss Medical Weekly, 142*, w13667.

[104] The clinical death in NDEs have all the relevant features of irreversible death to provide evidence supporting the claims of NDE advocates.

Advocates of Staunch Hylomorphism will favor a biological or organic criterion (aka the cardiopulmonary view) according to which there is an irreversible loss of biological processes or functions of the organism as a whole. This criterion needs further clarification. It does not entail that if a specific organ, system, or process has irreversibly failed (and may be "sustained" by external devises), the organism is dead. Rather, if there is evidence that the organism is still capable of exhibiting immanent causation, of sustaining its own functioning independently of external devices, then death has not occurred. While most today employ the brain position, often under the pressure to harvest organs for transplants, the debate is still a live one. We acknowledge that both are SD positions, but we suggest that Staunch Hylomorphists do more interdisciplinary work in defending and teasing out the implications of their ontology.

Another important issue is what it means for a human person to be healthy or, say, to have a healthy organ such as a heart. Though physicalism does not entail that humans lack essences, nevertheless, those physicalists who adopt an evolutionary, naturalist account of human origins (as we have argued in Chapter 2, this approach seems to imply strict physicalism) have an arduous task if they try to embrace essentialism. The two make very odd bedfellows. Accordingly, the preeminent philosopher of biology, David Hull, stated what he took to be obvious:

> The implications of moving species from the metaphysical category that can appropriately be characterized in terms of 'natures' to a category for which such characterizations are inappropriate are extensive and fundamental. If species evolve in anything like the way that Darwin thought they did, then they cannot possibly have the sort of natures that traditional philosophers claimed they did. If species in general lack natures, then so does *Homo sapiens* as a biological species. If *Homo sapiens* lack a nature, then no reference to biology can be made to support one's claims about 'human nature.' Perhaps all people are 'persons,' share the same 'personhood,' etc., but such claims must be explicated and defended *with no reference to biology*. Because so many moral, ethical, and political theories depend on some notion or other of human nature, Darwin's theory brought into question all these theories. The implications are not entailments. One can always dissociate '*Homo sapiens*' from 'human being,' but the result is a much less plausible position.[105]

Similarly, E. Mayr confidently declares:

> The concept of unchanging essences and of complete discontinuities between every eidos (type) and all others make genuine evolutionary thinking impossible. I agree with those who claim that the essentialist philosophies of Aristotle and Plato are incompatible with evolutionary thinking.[106]

If Hull and Mayr are correct, then it seems to us that evolutionary naturalists/strict physicalists have two options regarding how to understand, say, a healthy heart, both of which are strictly descriptive: Such a heart is either one that lies within a certain range on a bell curve or it is one that, in fact, confers an evolutionary advantage on its possessor.[107]

---

[105] Hull, D. (1989). *The metaphysics of evolution* (pp. 74–75). Albany, NY: State University of New York Press.

[106] Mayr, E. (1970). *Populations, species, and evolution* (p. 4). Cambridge, MA: Harvard University Press.

Substance dualists are in no such mess. SDists hold that *qua* substance, the embodied soul is an essentially characterized particular. Thus, a healthy heart is normative—one that functions the way it ought to function according to the nature of the organism that has it. We urge substance dualists to conduct research in this line of thinking to see if they can strengthen this argument and make it more precise.

Turning to mental health, studies show that widespread belief in SD continues to influence mental health and medical sciences.[108] Numerous studies have linked specific behaviors toward healthcare and mental health to the belief that SD is true. Some studies claim that belief in SD may increase stigma against mental health issues by weakening neurochemical explanations of mental disorders,[109] even leading some clinicians to perceive medication as less effective and therefore used less often.[110] However, adopting physicalism in medical school is correlated with prioritizing diseases or sub-personal analysis over the patient or holistic, personal analysis.[111] And although SD belief is blamed for a rise in individualism, studies have shown the opposite.[112] Although these studies have important implications, they are underdeveloped. This is an area for collaborative work by SDists across relevant disciplines.

11.4.6 SD Approaches to Developing Technology

Just as SD has implications for environmental ethics (e.g., to what extent am I a part of nature, and does that matter for how I ought to relate to nature and natural-things?), we should expect SD to impact the ethics of virtual and augmented environments, and the philosophy of technology in general.

---

[107] Alvin Plantinga makes a similar point regarding epistemology, warrant, and justification. See his *Warrant and proper function* (NY: Oxford University Press, 1993), 45–46, 194, 199–215. He does apply his claims to the notion of health. See 200, note 10.

[108] For example, a 2006 study concluded that mental health professionals continue to employ SD when analyzing clinical cases. A 2009 study concluded that “more than one-third of medical and paramedical professionals regarded mind and brain as separate entities” and that “the widespread dualism revealed by our survey continues to exert an influence on scientific thought.” See, Miresco, M., & Kirmayer, L. (2006). The persistence of mind–brain dualism in psychiatric reasoning about clinical scenarios. *American Journal of Psychiatry*, *163*(5), 913–918; and Demertzi, A., et al. (2009). Dualism persists in the science of mind. *Disorders of Consciousness: Annals of the New York Academy of Sciences*, *1157*, 1–9.

[109] Cf., Ventriglio, A., & Bhugra, D. (2015). Descartes’ dogma and damage to western psychiatry. *Epidemiology and Psychiatric Sciences*, (5), 368–370; Lebowitz, M. S. (2014). Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice*, *21*, 67–83; and Miresco, M. J., & Kirmayer, L. J. (2006). The persistence of mind–brain dualism in psychiatric reasoning about clinical scenarios. *American Journal of Psychiatry*, *163*, 913–918.

[110] Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians’ beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, *33*, 147–182.

[111] Whatley, S. D. (2014). Borrowed philosophy: Bedside physicalism and the need for a sui generis metaphysic of medicine. *Journal of Evaluation in Clinical Practice*, *6*, 961–974.

[112] Forstmann, M., & Burgmer, P. (2017). Antecedents, manifestations, and consequences of belief in mind–body dualism. In C. M. Zedelius, et al. (Eds.), *The science of lay theories* (pp. 189–190). Springer; The authors cite Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*, 382–386; and Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition, 13*, 103–128.

11.4.6.1 SD and Transhumanism

According to transhumanism, current human nature can be improved by the use of technological advancements, which may make it possible to increase human health, extend the human lifespan, and extend our intellectual and physical capacities. Technologies include genetic engineering, information technology, and anticipated future technologies such as fully immersive virtual reality, machine-phase nanotechnology, and artificial intelligence.[113] Most are also committed to *extreme rational essentialism*, the thesis according to which reason/cognition (reducible to information manipulation or computation) is at the core of human nature, such that posthumans will emerge from a suitably advanced cognitive capacity and a radically diminished negative affective capacity.[114]

The relationship between SD and transhumanism is underexplored. Plausibly, SD is incompatible with transhumanism in at least two ways. First, SD is not committed to and is likely incompatible with extreme rational essentialism. Computationalism about rationality/cognition is hard to square with a simple immaterial substance. And SDists have not reduced the essence of human persons to the capacity for reason alone. Even Descartes's notion of thinking in "thinking things" includes emotions.

Second, if SD is true, then plausibly, human nature cannot be changed at all, much less into a posthuman nature. The transhumanist views the body as that which, when altered in the right way alone, leads to human flourishing, culminating in posthumans.

The first element of a transhumanist moral vision is that the effort to address the human condition requires that we change the physical facts that, in part, generate the human condition. Curing the human condition requires altering the "human" part of the equation.[115]

The transhumanist view of the body is obviously incompatible with at least some versions of SD, such as Integrative SD and Staunch Hylomorphic SD. Facts about the body are grounded in facts about the soul, including human nature. However, things might not be so simple. We urge substance dualists to take on work that furthers an SD understanding of the limits of human enhancements and the implications that might follow from implementing transhumanist programs.

11.4.6.2 SD and Artificial Intelligence

Eric Olson rightly argues that AI research focuses on artificial thinking (e.g., a program) while failing to realize the importance of the ontology of artificial thinkers (e.g., the hardware).[116] Susan Schneider, a leading philosopher of AI, argues at length that AI research has failed to, but must, focus on their metaphysics of the subject of consciousness, as the most significant ethical implications of AI depend on what an artificial subject could or couldn't be.[117]

---

[113] This is a slightly modified definition from one of the most influential documents on transhumanism: Humanity+, "Transhumanism FAQ 3.0," Online Document (2020). https://www.humanityplus.org/transhumanist-faq.

[114] While those such as Nick Bostrom see transhumanists' extreme rational essentialism as compatible with Aristotle's rational essentialism (e.g., *De Anima*, II.3), Susan B. Levin right points out that Aristotle holds to the necessary role of nonrational faculties and intrapsychic harmony. "It follows that, by Aristotle's lights," says Levin, "embracing transhumanists' commitment to maximal augmentation of rationality per se as the sole rational course would testify to one's irrationality." Levin, S. B. (2021). *Posthuman bliss? The failed promise of transhumanism* (p. 4). Oxford: Oxford University Press; see, also, 68–84. See Bostrom, N. (2008). Why I want to be a posthuman when I grow up. In B. Gordijn & R. Chadwick (Eds.), *Medical enhancement and posthumanity* (p. 130). Dordrecht: Springer.

[115] Hopkins, P. (2008). A moral vision for transhumanism. *Journal of Evolution and Technology*, *19*(1), 4.

It is often argued that SD was written into the AI project. Some argue this began with a basic premise of the Human Genome Project, that information is separable from its material substrate. Algorithms are separable from and act on data. But the relationships between SD and AI are not so clear.

As of now, SDists have yet to enter this discussion, although one of us (Rickabaugh) is currently the principal investigator of a multi-year research project on this subject.[118] Schneider briefly explores SD throughout her book.[119] Although she rejects SD, she acknowledges that,

> Although you personally may reject the view that there's an immaterial mind or soul, science alone cannot rule it out … Contemporary substance dualists offer sophisticated nontheistic positions, as well as intriguing and equally sophisticated theistic ones.[120]

Schneider's acknowledgment reveals at least two things. First, SD is and rightly should be acknowledged as a sophisticated philosophical view. Second, SD is relevant to AI research.

Consider mind-uploading or whole brain emulation. Anders Sandberg and Nick Bostrom distinguish between[121]

> *Mind emulation*: A brain emulator that is detailed and correct enough to produce the phenomenological effects of a mind.
>
> *Person emulation*: A mind emulation that emulates a particular mind.

The Transhumanist eschatology is that of transferring a person—a conscious mind—from a biological substratum (the brain and central nervous system) to an artificial substratum (a computer), which can then be placed in a robotic or virtual body or perhaps a new biological substrate. It will be by person emulation, not mind emulation, that I am uploaded just prior to the death of my bio-body so that I can persist in an artificial afterlife.

But is person emulation possible on SD? It is not easy to see how personal emulation does not assume that a person is not a substance (immaterial or otherwise), but an aggregate of information ordered in some specific way. That is what is transferred from a bio-body to an artificial body. But this runs counter to the common charge that transhumanism assumes SD. SDists have much to explore here.

---

[116] Olson, E. (2019). The metaphysics of artificial intelligence. In M. P. Guta (Ed.), *Consciousness and the ontology of properties* (pp. 67–84). New York, NY: Routledge.

[117] Schneider, S. (2019). *Artificial you: AI and the future of your mind*. Princeton, NJ: Princeton University Press.

[118] An excellent exception is Cutter, B. (forthcoming). The AI ensoulment hypothesis. *Faith and Philosophy*.

[119] Schneider (2019), e.g., 76–77, 88, 128–129, 140–141.

[120] Ibid., 128–129.

[121] Sandberg, A., & Bostrom, N. (2008). *Whole brain emulation: A roadmap*. Technical Report #2008-3, Future of Humanity Institute, Oxford University. www.fhi.ox.ac.uk/reports/2008-3.pdf, 7–8.

Here is a further project. Richard Swinburne distinguishes between the following.[122]

> *Categorical theory*: Beliefs and desires inhere in the soul, where they exist even when we are not conscious of them.
>
> *Dispositional theory*: Beliefs and desires are simply dispositions of the brain to produce conscious feelings of desire and belief (as well as inclinations to act in certain ways) when we think about what we desire or believe.

Whichever view is true has significant implications for transhumanist uploading and old-fashioned life after death. On the categorical view, I carry into my afterlife the beliefs and desires which I have formed in this life. However, on the dispositional view, all that could survive would be an immaterial particular without any character to it. God would have to provide my character for my re-embodiment. Regarding transhumanism, on the categorical view I carry into my uploaded life the beliefs and desires which I have formed in my pre-uploaded life. Yet, on the dispositional view, it seems that nothing would survive. What remains after uploading are not my desires or my beliefs. Those are dispositions of my brain, a brain that no longer exists. What would it mean for the dispositions of my brain to persist without my brain persisting?

Here are some further lines of inquiry. How is information storage divided between the soul and brain? For example, when you become more virtuous, does your soul change or your brain or both? Suppose a virtuous person got a total data transfer from the brain of a vicious person. Would they become vicious?[123]

11.4.6.3 SD and Virtual Reality

The ongoing advancement of virtual reality and augmented reality raises several questions about the nature and phenomenology of virtual embodiment. It seems to some that cyberspace is the ultimate environment for Cartesian souls.[124] But is this true of all SD views? And what exactly can SDists say about the virtual body? Does a virtual body satisfy SD conditions for individuating one's body (e.g., direct causal interaction)? Is a virtual body, on SD, a cultural or social construction? Do the virtues of embodiment apply to virtual bodies as well as biological bodies? Can one person have two bodies, one biological and one virtual? Are there unique SD conditions for a virtual body to be a representation of a biological body or of the biologically embodied person?

We've mentioned only a few research projects here and urge others to develop these and other projects that employ unique SD insights and approaches to developing technology, especially VR and AI.

---

[122] See, Swinburne, R. (2013). *Mind, brain, and free will* (pp. 85, 167–169). Oxford: Oxford University Press.

[123] We are grateful to Richard Swinburne and Alexander Pruss for some of these suggested projects.

[124] See, e.g., Ajana, B. (2005). Disembodiment and cyberspace: A phenomenological approach. *Electronic Journal of Sociology*, *7*(2): 1–10. Cf., Young, G., & Whitty, M. T. (2010). In search of the Cartesian self: An examination of disembodiment within 21st-century communication. *Theory and Psychology*, *20*(2): 209–229.

11.5 CONCLUSION

In this chapter, we have proffered a number of important distinctions among SDists, along with suggesting several research projects for SDists. The reader may disagree with several of these projects, and some will be more controversial than others. Regardless of that, one thing seems clear: SD provides resources for numerous creative and likely fruitful research projects for the 21st century.

Appendix: The Soul, Mental Action, and the Conservation Laws

Mihretu P. Guta

In what follows, I will respond to three interrelated but distinct questions that collectively focus on whether the soul exerts causal influences upon the physical states or activities of the brain. Here are the three questions:

Q1: If the soul is constantly acting upon the brain, then why don't we see physically uncaused spikes in the energy level of the brain?

Q2: Are the neurons in the brain sufficiently sensitive to respond to such tiny stimuli as would be within the range of energy fluctuations permitted by the Uncertainty Principle? Is the brain responsive to quantum stimuli?

Q3: Would it be possible to show that the soul can affect the brain without our observing mysterious energy variations?

My discussion of questions (Q1)–(Q3) aims to reinforce the detailed responses given in Chapter 10 to some of the key objections the skeptics raise against substance dualism (SD). That said, the skeptics and the non-skeptics alike can raise questions (Q1)–(Q3) concerning the ontology of the soul. But for present purposes, I will discuss these questions primarily from the standpoint of those who embrace (in some form or another) the ontology of SD but still may wonder if the causal role of the soul or mental action can be empirically detectable. Broadly speaking, seeking an empirical basis concerning the causal power of the soul over the brain's operation arises from what I call the *Human*

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

Unobservability Concerns (HUC).[1] Unlike the brain's electro-physio-chemical processes, which are said to be directly observable, say, through imaging techniques, the soul is not observable in the same way. Despite the soul's unobservability, however, I will argue that there are excellent reasons to think that the reality of its causal power over brain activities remains intact.

Of course, the interaction(s) between the soul and the brain, arguably despite being contingent, is not one-directional. That is, what happens to the brain physiologically does affect what happens to the soul psychologically and vice versa. For example, when one is anxious or depressed, one's experience has both physiological and psychological dimensions.[2] Given that there is a two-way interaction between the soul and the body, figuring out the nature of the interaction in question lies at the heart of the nonempirical and empirical inquiries.[3] However, contra currently dominant trends in some philosophical and scientific circles, the empirical inquiry does not enjoy unconditional superiority over nonempirical inquiry. As we shall see, a two-way dynamic interaction between the soul and the body is fundamentally a metaphysical/ontological issue. That said, we can readily admit that we still lack complete knowledge of the nature of the interaction in question. Hence, the *mind–body problem*.[4] In light of this, I now turn to my discussion of the three questions introduced earlier.

---

[1] In his *A treatise of human nature*, Hume famously remarks, "For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception. When my perceptions are removed for any time, as by sound sleep, so long am I insensible of myself, and may truly be said not to exist. And were all my perceptions removed by death, and could I neither think, nor feel, nor see, nor love, nor hate, after the dissolution of my body, I should be entirely annihilated, nor do I conceive what is further requisite to make me a perfect nonentity." Hume, D. (1964). *A treatise of human nature* (p. 239). London: Dent. Hume mistakenly extrapolated from the invisibility or unobservability of the self (soul) the conclusion that no such thing as the self exists in the first place. Ironically enough, Hume's unobservability criterion, which he relies on to deny the reality of the self, unavoidably presupposes the reality of the very thing he so strongly dismisses. I do not have space to extend this discussion any further. But the point should be clear from a closer study of the passage quoted above.

[2] For some interesting empirical evidence on this matter, see Amen, D. G. (1998). *Change your brain change your life: The breakthrough program for conquering anxiety, depression, obsessiveness, anger, and impulsiveness*. New York, NY: Three Rivers Press.

[3] Elsewhere I tried to show how these two modes of inquiries can feed into each other, see Guta, M. P. (2019). Introduction. In M. P. Guta (Ed.), *Consciousness and the ontology of properties* (pp. 1–12). New York, NY: Routledge.

[4] For an excellent discussion on the role of metaphysics/ontology in setting the tone for empirical investigation of the nature of reality, see Lowe, E. J. (2006). *The four-category ontology: A metaphysical foundation for natural science*. Oxford: Oxford University Press; and Lowe, E. J. (2002). *Introduction to the philosophy of mind*. Cambridge: Cambridge University Press. Also see, Moreland, J. P. (2018). The fundamental limitations of cognitive neuroscience for stating and solving the ubiquitous metaphysical issues in philosophy of mind. *Philosophia Christi*, *20*(1), 43–51; and Guta, M. P. (2021). Metaphysics, natural science and theological claims: E. J. Lowe's approach. *TheoLogica: An International Journal for Philosophy of Religion and Philosophical Theology*, *5*(2), 129–160.

Q1: *If the soul is constantly acting upon the brain, then why don't we see physically uncaused spikes in the energy level of the brain?*

Two Conditions

As I see it, some progress can be made in answering (Q1) if there are ways to satisfy, at least, two necessary conditions, namely:

(a) The soul and its actions must be kinds of things that are directly observable in the sense of being empirically identifiable or distinguishable from the brain's complex neurophysiological processes/events of which the electrochemical signals are central.
(b) There must be some way by which we come to determine whether we have succeeded in establishing (a).

I take it to be the case that (a) is an ontological condition since satisfying it requires establishing first what sorts of things the soul and its actions are; whereas (b) is an epistemological condition since it focuses on the criterion or method by which something is done. In this case, the "something" in question consists in demonstrating how (a) can be satisfied. But dealing with (a) and (b) would require one to grant (at the minimum) the *possibility* of the existence of the soul and its causal influence upon the brain. Of course, here, the possibility in question could take different forms, say, *conceptual*, *epistemic*, *logical*, *metaphysical*, or *nomological*.

For example, X is said to be conceptually possible, if X's possibility can be conceptually entertained. X is said to be epistemically possible, if, for all one knows, X is the case. X is said to be logically possible if X's possibility is not logically contradictory. Some philosophers think that logical possibility is almost identical to metaphysical possibility. But the metaphysical possibility can be captured in its broadest sense. For example, it is not physically possible in the actual world, for any human being to travel faster than the speed of light. Yet, in some other world with different sets of laws, it could be metaphysically possible for human beings to transcend the current universal limit of physical speed, that is, *c*, which equals 2.998×10^{10} cm sec^{-1}. X is said to be nomologically possible if X's possibility does not violate the laws of nature.[5]

Assuming the actual metaphysical possibility of the soul and also its causal influence upon the operations of the brain, (Q1) mainly requires us to figure out if there are (could be) physically uncaused spikes in the energy level of the brain. There is an immediate problem we face in tackling this issue. It is not inaccurate to say that people who work in the fields of cognitive science, cognitive neuroscience, cognitive psychology, and neurology do not carry out their research with the goal in mind to see the role of the soul and its causal impact on brain function/activity. That means that people who work in these fields do not deal with (Q1) at all; not at least, in any direct or intentional manner. This means that we can't just assume or expect the empirical works carried out in these fields to give us insights into answering (Q1).

---

[5] See further Gendler, T. S., & Hawthorne, J. (2002). *Conceivability and possibility*. Oxford: Oxford University Press; Jubien, M. (1997). *Contemporary metaphysics*. Oxford: Blackwell Publishers, chapter 8; Conee, E., & Sider, T. (2014). *Riddles of existence: A guided tour to metaphysics*. Oxford: Oxford University Press, chapter 9; Lowe, E.J. (2002). *Survey of metaphysics*. Oxford: Oxford University Press, chapters 5–7; and Hoffman, J., & Rosenkrantz, G. S. (1994). *Substance among other categories*. Cambridge: Cambridge University Press, chapter 5.

Belief in the actual existence of the soul is a minority position among philosophers and scientists. The notion of the soul is being sidelined by practicing neuroscientists. For example, a Vanderbilt University award-winning neuroscientist, Jeanette Norden argues that it is not for neuroscience to worry about questions related to the existence of the soul. In Nordan's view, neuroscience does not and cannot tell us whether there is some kind of quality [soul] present in humans that will survive the death of the brain and the death of the body. As Nordan sees it, modern neuroscience aims to show and understand how the brain and, ultimately, the mind (i.e., some kind of higher-order brain property) are related.[6]

Similarly, in his *The Philosophy of Cognitive Science*, M. J. Cain remarks, "most mainstream cognitive science is underwritten by a commitment to the anti-dualist view that the mind is in some substantial respect the brain, so the study of cognition is the study of the workings of a physical system."[7] In any case, however, let us assume that people working in the scientific fields mentioned above want to tackle (Q1). In that case, how might they go about doing just that?

Scanning Brains

From an empirical standpoint, the most promising way to answer (Q1) would be to scan or peer into a living human brain using the appropriate brain imaging techniques. In this case, the brain imaging techniques we need are not the ones that are being used to image the structure or anatomy of the brain, namely *computerized tomography* (CT) and *magnetic resonance imaging* (MRI). Such scanning machines play an important role in cancer screening *inter alia*. But they cannot be used to correlate a certain cognitive activity, say, Mary's attempt to solve a linear algebraic equation with brain activity.

The imaging techniques that are often employed to study cognition and its relation to brain activity include *functional magnetic resonance imaging* (fMRI), *positron emission tomography* (PET), and an *electroencephalogram* (EEG). For example, fMRI allows us, among other things, to detect cerebral blood flow, metabolic activity, and BOLD (i.e., blood oxygenation level dependence). In the regions of the brain where blood flow is higher, it is said that there will always be corresponding neuronal firings. Where cerebral blood flow increases, the brain uses up more energy, thereby facilitating a person's engagement in some mental activity. Similarly, PET allows us to detect blood flow in brain regions. In this case, a safe amount of radioactive tracers are given via intravenous (IV) injection to locate the electrical signals. EEG uses electrodes on the skull to record electric impulses. This technique allows us to assess brain activities by analyzing brain waves.[8] Cognitive neuroscientists attempt to correlate brain activity with mental tasks based on physiological data obtained via such brain imaging techniques.

---

[6] See for details, Nordan, J. (2007). *Understanding the brain*. Lecture Series, Parts 1–3 (p. 6). Virginia: The Teaching Company. See also a similar remark by Springer, S. P., & Deutsch, G. (1981/1998). *Left brain right brain: Perspectives from cognitive neuroscience* (p. 333). W.H. Freeman and Company, Worth Publishers.

[7] Cain, M. J. (2016). *The philosophy of cognitive science* (p. 7). Cambridge: Polity Press. See also similar works that do not deal with the ontology of the soul: Bickle, J. (Ed.). (2009). *The Oxford handbook of philosophy and neuroscience*. Oxford: Oxford University Press; Cummins, R., & Cummins, D. (Eds.). (2000). *Minds, brains, and computer: The foundations of cognitive science*. Oxford: Blackwell Pub; and Stainton, R. J. (Ed.). (2006). *Contemporary Cognitive Science*. Oxford: Blackwell.

Despite some of the crucial roles that they play, the aforementioned imaging techniques have limitations. In this regard, Heisenberg's Uncertainty Principle can illustrate the limitations in question. Mathematically this principle is expressed as follows:

$$\Delta x \Delta p \geq \hbar / 4\pi$$

In this equation, the delta sign refers to *uncertainty*, and ħ is Planck's constant, which is an extremely small number, 6.626×10^{-34} J.s. According to *Heisenberg's Uncertainty Principle*, a simultaneous specification of the precise position of a particle, Δx, and the precise momentum, Δp, of a subatomic particle is impossible. That is, if we get the precise information regarding the position of a particle, Δx, then we lose the precise information regarding its momentum (i.e., mass multiplied by velocity, $p = mv$) and vice versa.[9]

Similarly, in the case of neuronal firings, establishing complete information regarding the temporal and spatial resolutions of brain imaging results proves to be difficult.[10] Here the term "temporal resolution" refers to the time-related specification of neuronal activities whereas the term "spatial resolution" refers to the spatial location-related specification of neuronal activities. For example, if we get information about spatial resolution, then we will lose information about temporal resolution and vice versa. Somewhat superficially mimicking, I extend the standard equation used to describe quantum uncertainty to that of the situation we face in brain imaging. In this case, Δ stands for uncertainty, tr stands for *temporal resolution* and sr stands for *spatial resolution*. Putting these things together, we can then suggest an applied formal equation in the following way: $\Delta tr \cdot \Delta sr \geq \hbar/4\,\pi$. This equation reads: the product of temporal resolution and spatial resolution is greater than or equal to Planck's constant divided by 4 π. I leave the mathematical details for another occasion.

---

[8] For a very useful and detailed analysis on brain imaging techniques, see Bechtel, W., Mandik, P., et al. (Eds.). (2001). *Philosophy and the neurosciences: A reader*. Oxford: Blackwell, Parts I-V. Research on the human brain has come a long way. Until recently, the main method used to study the human brain has been either dissecting of animals' brains: or carrying out work on deceased human bodies. But brain imaging techniques have opened up new effective ways. But this does not mean that researchers are totally out of the woods yet. In fact, researchers still depend heavily on animal researches and human autopsies to precisely understand various sorts of neurological issues.

[9] Planck's constant is said to determine the discrete unit of energy, mass, spin, and similar other quantum phenomena. In this case, Planck's constant sets a fundamental parameter in quantum mechanics. Quantum uncertainty does not arise as a result of an immature technology. Rather, reality is such that no complete information can be had regarding these two quantities at the same time.

[10] It could be said that no similar uncertainty phenomenon can be extended to the limitations that the brain imaging techniques pose to us when it comes to acquiring complete information about the temporal and the spatial resolution. As technology improves, this problem could be said to have a lasting solution. However, I am inclined to suspend my judgment on this issue. What reasons do we have to think that in the case of brain imaging techniques our challenge is merely technological? I don't think we can prejudge this situation. It may well be the case that the way the brain functions is such that we cannot be in a position to get information regarding the temporal and the spatial resolutions at the same time. If this turns out to be the case, then nothing prevents us from attempting to extend the case of uncertainty to the situation we face in brain imaging scenarios as well. I will illustrate what that looks like in what follows.

Ideally speaking, when neuronal signals are recorded, it will be advantageous to get simultaneously complete information regarding the time it takes for a certain relevant population of neurons to fire as well as where in the brain those electrical signals originated. As of yet, no techniques are available that would allow us to gather both temporal and spatial information in one experimental setting to a satisfactory degree. As things stand, certainty about Δtr means less certainty about Δsr and vice versa. Let us call this the *Phenomenon of the Uncertainty of the Spikes Problem*.

For example, EEG has good Δtr but poor Δsr. PET has good Δsr (in the order of millimeters) but poor Δtr. fMRI has good Δsr but poor Δtr. Even if techniques that would allow us to establish simultaneously both Δtr and Δsr were available, our task would not be easy given the sheer quantity of neurons (approx. 100 billion) and neuronal synaptic connections (approx.100 trillion) that we would have to deal with. To this list we can add chemical signals that operate in highly complex ways in conjunction with electrical signals (not to mention other physiological processes). Moreover, taken in themselves, the brain's electrochemical signals mean nothing until they are interpreted. But interpreting the signals in question gets very tricky when the intent is to correlate them to some mental content. Interpretations are often susceptible to flawed inferences (not to mention the impact of problematic methodological models) regarding what the signals mean.[11]

Of course, imaging techniques are good tools to scan the physical brain but not the mental life of people or identify the soul and its actions. This is because, as I argue elsewhere, human beings have irreducible first-person access to their own subjective experiences.[12] That means that accessing someone's subjective experiences via using imaging techniques is a nonstarter. In this case, I fully agree with an eminent neuroscientist, Jeffrey M. Schwartz, and Sharon Begley's observation:

> Not even the most detailed fMRI gives us more than the physical basis of perception or awareness; it doesn't come close to explaining what it feels like from the inside. It doesn't explain the first person feeling of red. How do we know that it is the same for different people? And why would studying brain mechanisms, even down to the molecular level, ever provide an answer to these questions?[13]

---

[11] See for example, Bechtel, W., Mandik, P., et al. (Eds.). (2001). *Philosophy and the neurosciences: A reader* (pp. 55-57ff). UK: Blackwell. Although I disagree with their analysis of the phenomenon of pain, V. G. Hardcastle & C. M. Stewart point out some serious problems regarding fMRI technique. See their (2009). fMRI: A modern cerebrascope? The case of pain. In J. Pickle (Ed.), *The Oxford handbook of philosophy and neuroscience*. Oxford: Oxford University Press, chapter 8. Despite my sharp disagreement with their conception of the role of philosophy which they relegated to making mere clarifications of concepts, M. R. Bennett and P. M. S. Hacker also discuss a wide range of problems in empirical works; see *Philosophical foundations of neuroscience*. (2003). Oxford: Blackwell.

[12] See e.g., Guta, M. P. (2015). Consciousness, first-person perspective, and neuroimaging. *Journal of Consciousness Studies*, *22*(11–12): 218–245. In my forthcoming two papers entitled, "Mirror neurons, consciousness and the bearer question" and "In what sense is consciousness a property," I make positive cases as to why empirical works in cognitive neurosciences need a metaphysical foundation to have any hope of being informative and relevant.

[13] Schwartz, J. M., & Begle, S. (2002). *The mind and the brain: Neuroplasticity and the power of mental force* (p. 27). New York: HarperCollins.

In light of Schwartz's and Begle's remarks, I too do not see any promising way to satisfy (a) via imaging techniques. If I am right about this, then it follows that it is altogether an open matter as to what constitutes satisfying (b) either. Recall that (a) pertains to a necessary condition regarding the observability of the soul and its actions on brain functions, and (b) pertains to a necessary condition of establishing ways (criteria) that are used to determine if (a) is satisfied.

The Entropic Brain Hypothesis

That said, however, imaging techniques such as fMRI are said to play a key role in helping us understand the energy level of the brain. In this case, currently, the most popular model is known as *The Entropic Brain*. This model was first introduced by Robin L. Carhart-Harris. In a paper published in 2014 entitled "The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs," Carhart-Harris and his collaborators introduced the theory of consciousness that incorporates the principles of physics, neurobiology, and psychoanalysis.[14] This study mainly focuses on tackling questions related to how the normal waking consciousness of healthy adult humans is related to different kinds of states of consciousness. In this case, Carhart-Harris and his collaborators looked at how the brain maintains its normal state of waking consciousness. They also looked at what happens to the brain function in the case of non-ordinary states such as rapid eye movement (REM), sleep and dreaming, early psychosis, and the psychedelic state of consciousness.

Given the entropic brain hypothesis, it is said that the quality of any conscious state depends on the system's entropy which is measured against the backdrop of the state of the brain function. Carhart-Harris and his collaborators characterized entropy as the dimensionless quantity used to measure uncertainty in a given system. In this case, entropy is taken as a physical quantity. High entropy is said to be a form of high disorder in a system. But how does entropy work in the case of the brain?

To show how entropy works in this case, Carhart-Harris and his collaborators made a distinction between *primary consciousness* and *secondary consciousness*. The primary consciousness is said to consist of dreaming, onset-phase psychosis, near-death experience, the sensory-deprived state, and the psychedelic drug state. In the primary state, entropy is said to be high due to the disruption of self-organized brain function via say, psychedelic drugs. In a primary state of consciousness, people are said to come up with wishful beliefs in which they attempt to link their experience(s) to some sort of supernatural base. In this case, belief in the reality of the soul and its action on the brain function would be taken as an expression of a wishful belief/thinking.

By contrast, Carhart-Harris and his collaborators argue that there is always low entropy in the case of secondary consciousness. This sort of consciousness pertains to a normal waking state. In this case, the brain is said to suppress entropy thereby blocking its increase. It does this by keeping self-organized brain function to stay stable so that metacognition (the second-order ability for self-reflection) would continue intact. But I must point out that the author's characterization of the brain engaging in self-organized mental activity is problematic. For example, Bennett and Hacker, in their book entitled *Philosophical Foundations of Neuroscience*, deem attributing psychological capacities to the brain as a mereological fallacy;[15] that is, attributing something to the part which in reality belongs to the whole.

---

[14] See Carhart-Harris, R. L., et al. (2014). The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in Human Neuroscience*, *8*(2), 1–22.

Here Carhart-Harris and his collaborators are making familiar mistakes often made in most empirical works in which the brain is being portrayed as an agent. That is, "the brain says this and that," "the brain refuses," "the brain thinks," and so on. In any case, this study goes into a lot of detail in terms of showing the benefit of psychedelic drugs to understand the state of the brain in its high entropy state. For reasons I cannot go into here, however, the method employed in this study to spell out the nature of mental states has several questionable assumptions. But by way of helping us tackle (Q1), this study offers us no guidance whatsoever. So both the ontological condition (a) and the epistemic condition (b) remain unsatisfied.

Alternative Proposal

Where should we go from here then? I do think that (Q1) cannot be handled adequately and fruitfully by engaging in a mere empirical study. So we need to engage in metaphysical reflection. In this case, we should guard against Humean blindspots, that is, assuming the physical observability of the soul and its actions on the brain function as a necessary precondition before affirming whether such features of reality are real. In other words, HUC, i.e., the *Humean Unobservability Concerns*, can be put aside. In doing so, we can free ourselves up from assuming any burden of proof to satisfy (a) and (b). Alternatively, I suggest E. J. Lowe's model of mental causation. Lowe's model provides us with an excellent framework to think why the soul and its actions upon the brain could be real but invisible thereby also being inaccessible for direct scientific manipulations.[16]

Lowe imagines a physical event *P* occurring at a time t_2 and a nonphysical or mental cause of *P* occurring at the earlier time t_1. Lowe argues that consistent with the *causal closure principle* which states that every physical event has only a physical cause, it is possible to represent *P* as having a complete physical cause at $t_{1,}$ the conjunction of physical events, $P_{11,}$ and P_{12}. Lowe argues that there is also room to represent *M* (mental cause) as having a physical event P_{12} as a sufficient cause. Lowe argues that since P_{12} occurs at the same time as that of *M*, its causation of *M* constitutes a case of simultaneous causation. In this case, Lowe argues that the non-overdetermination principle, which prevents the presence of two distinct causal sources (i.e., mental and physical) for a given effect, is violated. In the imagined scenario, the violation in question consists of *P* having at t_1 both a mental causa *M* and an entirely distinct physical sufficient cause, which is a conjunction of P_{11} and P_{12}. Figure A.1 shows how Lowe sums up his model of mental causation, with he takes to be a dualistic counterexample to the non-overdetermination principle.[17]

As Lowe sees it, an important lesson can be drawn from his model of mental causation. Lowe argues that any scientist who examines the causal process depicted in the model

---

[15] See, Bennett and Hacker (2003, chapter 3).

[16] See Lowe, E. J. (2008). *Personal agency: The metaphysics of mind and action*. Oxford: Oxford University Press, Chapters 1–5. Lowe extends this discussion in great detail. However, for present purposes, I stick only to the core aspect of Lowe's theory of mental causation.

[17] See Lowe (2008, pp. 68–74).

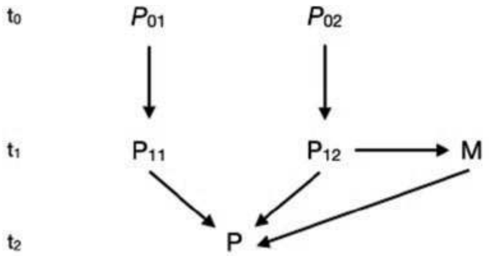


FIGURE A.1 Lowe's SD Mental Causation.

in question from an empirical standpoint most likely concludes that only physical causal relationships hold up. In this case, Lowe argues that the scientists in question may think that the physical event *P* had a complete and entirely physical causal explanation in terms of its immediate physical causes, namely P_{11} and P_{12}, which also includes its prior causes P_{01} and P_{02}. In such a situation, Lowe argues that our scientists wouldn't suspect the presence of any "gaps" in the causal sources before the occurrence of *P* at t_2.

Lowe argues that if the scientists in question were neuroscientists, it would be easy for them to think that *P* can be explicable entirely in terms of brain events or neuronal firings. However, Lowe argues that if the model of mental causation in question is right, our neuroscientists would be entirely mistaken because despite *P* having purely physical causes, it also has as its cause a nonphysical mental even *M*. But Lowe argues that the existence and the causal role of *M* would be *unobservable* to our neuroscientists given the causal structure of the situation depicted coupled with the limited empirical means that are available to undertake such an investigation.

In light of such considerations, Lowe further argues that not only would our neuroscientists not be able to locate *M*'s presence and its causal contribution directly via observation, but our neuroscientists would not be able to detect *M*'s presence indirectly either. Lowe thinks this would be true even in a situation in which our neuroscientists happen to realize that the causal account of *P* in entirely physical terms is incomplete. But as things stand, for our neuroscientists, the causal explanation of *P* in terms of physical terms appears to be complete. Lowe concludes, "we have no particular reason to suppose that this sort of situation is purely hypothetical: it may, for all we presently know, be the sort of situation that actually obtains with human minds and brains."[18]

Lowe's model of mental causation is a response to *the causal closure of the physical domain thesis*, that is, *every physical event has a sufficient physical cause*. In this case, nonphysical mental causes are ruled out from having any causal role in bringing about a physical event. Physicalists defend this thesis claiming that it perfectly obeys the constraints imposed by the conservation laws. According to these laws, a measurable quality or property of an isolated physical system does not change over time. Physicalist theorists of mind think that conservation laws prevent nonphysical causal input from playing any role in a closed physical system.

---

[18] Lowe (2008, p. 75).

Against this assumption, as we saw, Lowe introduced a sophisticated model of mental causation which shows that mental causation plays an integral role within a physical causal structure. Lowe shows how it is easy for scientists, who focus only on the physical aspect of what they study, to arrive at a mistaken conclusion regarding the role of mental causes. The mistake Lowe has in mind emerges due to the invisibility of mental causes. Scientists may think that mental causes are absent from the causal structure of the physical world. Lowe further shows that conservation laws are not even important in this debate. They play no practical or metaphysical role given that mental causation cannot be blocked from playing a causal role in a physical world. So a good case can be made for interactive dualism quite independently of what one thinks about conservation laws.

However, in his *The Engine of the Soul*, W. D. Hart argues that traditional Cartesian Dualism can be defended in a way that does not conflict with conservation laws. For example, Hart proposes “psychic energy” which he claims can be converted into physical energy thereby obeying the conservation laws.[19] But if as Hart argues, “psychic energy” exists, then energy is not necessarily a physical quantity. This has other ramifications for the causal closure of the physical thesis, but I postpone that discussion for another occasion.

A good case can be made to defend Lowe’s approach as well as Hart’s approach. But I have a long-standing skepticism when it comes to conservation laws. Take, for example, the *first law of thermodynamics*, which states that energy is neither *created* nor *destroyed*. Hardcore physicalists assume that this law provides them with an excellent empirical ground to establish their metaphysical claim, that is, *the causal closure of the physical domain*. But what exactly is the evidence for the *first law of thermodynamics* as providing reasons for us to think that the universe is a closed system? The central claim of this law seems to be more metaphysical than scientific. If this is right, then the law in question cannot serve as a piece of scientific evidence for another metaphysical claim, that is, the *closure principle*.

In my extensive interactions with various physicists, I was not able to get any clear and convincing answer on what ground we are supposed to take *the first law of thermodynamics* as a settled or confirmed hypothesis. This claim might be shocking to some but it need not be. For example, in their article entitled, “On the Status of Conservation Laws in Physics: Implications for Semiclassical Gravity,” Tim Maudlin, Elias Okon, and Daniel Sudarsky express similar but even more serious worries about the status of conservation laws in general. As they remark:

> If energy were nearly conserved, conserved in everyday life to within some very small quantity, the practical utility of these shortcuts would be just as great. But as logically compelling as this observation is, it does not address a deeper attachment to energy conservation as the fundamental physical principle of the universe. In a certain state of mind, even the tiniest, most practically insignificant violation of the conservation of energy is almost unthinkable. Attachment to the principle seems to run much deeper than mere empirical grounds can account for, as if the violation of the conservation law would be wrong.[20]

---

[19] Hart, W. D. (1988). *The engine of the soul*. Oxford: Oxford University Press.

[20] Maudlin, T., Okon, E., & Sudarsky, D. (2020). On the status of conservation laws in physics: Implications for semiclassical gravity, 2–3. https://core.ac.uk/download/pdf/295732981.pdf. Accessed 28 August 2022. The article was published in the (2020). *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *69*(February): 67–81.

Furthermore, Maudlin, Okon, and Sudarsky ask:

> This sense of the deep foundational status of the conservation of energy poses yet another question, viz. Why should energy be conserved? What is the source of this veneration of the principle, which makes any violation of it, however small, seem so momentous? This question leads us to the last in our series, the simplest and deepest question, which sheds light on all the rest, namely What is energy?[21]

Maudlin, Okon, and Sudarsky raised foundational questions. For lack of space, I cannot elaborate on them. But the issues raised require serious engagement. That said, it is one thing to say that in a localized controlled system one could hope to show that energy is conserved or not destroyed. But what exactly justifies the most serious claim that energy is something that was never created in the first place? In light of our impoverished knowledge of say, dark energy and dark matter,[22] there seems to be no good reason to think that the conservation laws, in general, provide us with any empirical ground to establish the *closure of the physical domain* thesis. Even if we had adequate knowledge of dark energy and dark matter, the question of whether the universe is causally closed remains to be something that requires a metaphysical basis, not just an empirical one.

In the absence of any compelling reasons to the contrary, nothing prevents us from taking seriously the claim that the universe accommodates mental causes as much as it does physical causes. The *first law of thermodynamics* suffers from other problems the details of which need not concern us.[23] Even if one embraces mental causes without upsetting conservation laws (Hart's approach) or ignores conservation laws entirely making an independent case for mental causation (Lowe's approach), the *first law of thermodynamics* does not seem to play any key empirical role in establishing the metaphysical claim: *the causal closure of the physical domain*.

Chalmers and McQueen

Chalmers and McQueen have co-authored a paper entitled, "Consciousness and the Collapse of the Wave Function."[24] The central aim of this paper is to explore if a theory can be established concerning the causal role of consciousness in collapsing the wave function.[25] The measurement problem remains unresolved. It is commonly assumed that the wave function collapse happens only if it involves measurement. But measurement as a precondition for quantum state collapses stands at odds with GWR-spontaneous collapse theory. In this theory, quantum state collapses are real but they are not triggered by any measurement. Collapses happen randomly. On the other side of the spectrum, Bohemian mechanics and Everettian many world theories rule out quantum state collapses altogether. Chalmers and McQueen explore an alternative way to understand the wave collapse, and in this case, they propose a consciousness-collapse view.

---

[21] Maudlin, Okon, and Sudarsky. 3.

[22] For example, theoretical physicist Lee Smolin remarks, "Recent measurements reveal a universe consisting mostly of the unknown. Fully 70 percent of the matter density appears to be in the form of dark energy. Twenty six percent is dark matter. Only 4 percent is ordinary matter. So less than 1 part in 20 is made out of matter we have observed experimentally or described in the standard model of particle physics. Of the other 96 percent, apart from the properties just mentioned, we know absolutely nothing," *The trouble with physics: The rise of string theory, the fall of a science and what comes next* (p. 16). (2006). New York: A Mariner Book.

[23] Here I have in mind the various cosmological models of the universe. The *first law of thermodynamics* can be shown to be inconsistent with certain cosmological models that presuppose the creation of matter.

[24] Chalmers and McQueen (2021, p. 3). https://arxiv.org/pdf/2105.02314.pdf. Accessed 20 August, 2022. This paper was published in an edited volume. See Chalmers, D. J., & McQueen, K. J. (2022). Consciousness and the collapse of the wave function. In S. Gao (Ed.), *Consciousness and quantum mechanics*. Oxford: Oxford University Press.

Chalmers and McQueen advance a highly detailed and mathematically informed discussion. Due to space limitations, I cannot go into any of the details of this great paper. But it is worth noting one relevant issue from their paper for the discussion at hand. If consciousness has a causal role in a physical world, then as we already argued earlier, the causal closure of the physical domain thesis will be seriously undermined. So those who embrace the closure principle deny the causal efficacy of consciousness. Chalmers and McQueen anticipate this very objection but find it wanting. As they remark:

> Interestingly, the most common reason among philosophers for rejecting property dualist theories of consciousness is an argument from physics. This argument runs roughly as follows: (1) every physical effect has only physical causes, (2) consciousness causes physical effects, so (3) consciousness is physical. The key first premise is a causal closure thesis, supported by the observation that there are no causal gaps in standard physics that a nonphysical consciousness might fill. But wave function collapse in quantum mechanics appears to be precisely such a gap, and consciousness-collapse models are at least not obviously ruled out by known physics.[26]

Furthermore, they remark:

> The situation is that many physicists rule out consciousness-collapse models for philosophical reasons (they are dualistic), while philosophers rule out property dualist models for physics-based reasons (they violate causal closure). The upshot

---

[25] A wave function is governed by Schrödinger equation. One of the mathematical formalizations of this equation looks like this:

$$i\hbar\,\partial\psi(x,t)/\partial t = \widehat{H}(\psi)(x,t)$$

The equation tells us that a given quantum state or wave function evolves/changes over time along the x-axis depending on the total energy of the system. So the equation tells us that a wave function in its superposed state evolves deterministically unless measurement is applied to it. The quantum system is represented by mathematical object, $|\psi\rangle$. When a measurement is applied to a quantum system, a definite state or an eigenstate value can be determined. In this case, $|\psi\rangle$ takes a certain value, say $|\psi_1\rangle$ corresponding to its energy level, say E_1 and so on. For details on theories of wave collapses, see the excellent discussion in Maudlin, T. (2019). *Philosophy of physics*: *Quantum theory*. NJ: Princeton University Press, chapters 3–6. For a more accessible discussion, see Gaymer, M. G. (2017). *Quantum physics*. Oxford: Oxford University Press.

[26] Chalmers and McQueen (2021, p. 8). https://arxiv.org/pdf/2105.02314.pdf.

> is that a central reason to reject the consciousness-collapse thesis (it leads to dualism) and a central reason to reject interactionist property dualism (it violates the causal closure of physics) provide no reason to reject the two views when taken together. Perhaps there are other reasons to reject the consciousness-collapse thesis or to reject dualism, but these reasons must be found elsewhere.[27]

In light of such considerations, I conclude that (Q1) cannot successfully be answered on an empirical basis alone.

Objection:

Of course, at this point, a follower of Jaegwon Kim might offer an alternative mental causation model to that of Lowe's (including Chalmers's and McQueen's) arguing that mental causes are not genuine. For example, Kim's mind/body supervenience grounds each mental phenomenon in the physical domain.[28] That is, a mental phenomenon has a set of physical conditions that are nomologically sufficient for it. However, mind–body supervenience itself is a source of the problem. For example, an instance of mental property M causes another mental property M* to be instantiated. In this case, M* has a physical supervenience base P*. But where does the instance of M* come from? According to Kim, two responses can be given: (i) M caused M*; (ii) P* the supervenience base of M*, is instantiated on this occasion.

But given the mind–body supervenience, M* occurs because of its supervenience base P* occurs, and as long as P* occurs, M* must occur regardless of whether or not an instance of M preceded it. So, the claim of M to be the cause of M* becomes shaky. If anything were to have a role in the causation of M*, it would be via its relationship to M*'s supervenience base P*. To claim M to be a cause of M* with the fact that M* has P* as its supervenience base is to accept that M caused M* by causing P*. This is a case of mental to physical causation.

If causation is taken as grounded in nomological sufficiency, P qualifies as a cause of P*. On the other hand, if we take causation counterfactually, again P qualifies. That is, if P had not occurred, M would not have occurred. But if we say that M had not occurred, P* would not have occurred. The same thing goes. That is, if P had not occurred, P* would not have occurred either. Hence, an overabundance of causes: M and P a sufficient cause of P*. But here we can take P to be the cause of P* and M supervenes on P and M* supervenes on P*. Hence, given the physicalists' account, the M-to-M* and M-to-P* causal relations are only apparent, arising out of a genuine causal process from P to P*.

Reply:

Kim's model of mental causation is highly problematic. Here I can only mention a few points. Elsewhere I made a case against it.[29] That said, the physicalists' case for mental-to-physical causal feedback is not plausible. Physicalism advocated by Kim is thoroughly reductionistic. Physicalists claim that if mental properties are taken as the nonphysical cause of physical events, then that would violate the causal closure of the physical domain. But this begs a question in that unless mental properties are reduced to a physical domain they would be irrelevant to establishing a causal relationship between the mental and the physical states. Why should we think this is the case? Since our mental states—our beliefs, desires, and intentions—have causal effects in the physical world, to deny that mental states have genuine causal powers is counterintuitive at the least.

---

[27] Chalmers and McQueen (2021, pp. 7–8). https://arxiv.org/pdf/2105.02314.pdf.

[28] See Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge: The MIT Press, chapters 1–2.

[29] See Guta, M. P. (2011). Frank Jackson's location problem and argument from the self. *Philosophia Christi*, *1*(13), 35–58.

Worse, physicalists claim that any physical event is not outside of the physical domain. Moreover, everything can be explained based on physics. Such is a strong philosophical claim which can hardly be accommodated within the physicalist domain. If this is right, to claim that a complete physical explanation of everything is possible is self-defeating. Furthermore, if mental states do not have genuine causal power, then they are merely epiphenomenal, which is hard to believe given what we know about intentional states and our own first-person awareness of them. Thus, no plausible physicalist case is made. It is hard to see the advantage of the Kimean model of mental causation over that of Lowe's. Given that Lowe's model aligns very well with our own experience, I take it to be superior to Kim's.

Q2. Are the neurons in the brain sufficiently sensitive to respond to such tiny stimuli as would be within the range of energy fluctuations permitted by the Uncertainty Principle? Is the brain responsive to quantum stimuli?

Settling the question of whether not the brain is sensitive to quantum stimuli depends to a large extent on our ability to distinguish normal brain activities from those of quantum events. As we shall see, drawing a line between these two sorts of phenomena might not be something that can be established in any straightforward manner. But before I voice out some of my concerns in this regard, in what follows, I briefly present an exposition of Peter Jedlicka's paper entitled, "Revisiting the Quantum Brain Hypothesis: Toward Quantum (Neuro)biology?"[30]

This paper argues that the brain is responsive to tiny quantum fluctuations. Jedlicka begins his discussion by pointing out that viewing the nervous system as a linear, computer-like machine that functions in a deterministic classical input-output model of computation is still popular in neuroscience.

However, Jedlicka claims that recent experimental and theoretical works show that the popular model is unsatisfactory. Instead, the nervous system is a nonlinear dynamical system displaying stochastic activity. Neural networks are made up of nerve cells that are connected in numerous reciprocal ways. The interwoven network is what is said to allow the brain to have a capacity to engage in nonlinear computations thereby also generating dynamic feedback loops. In light of this, it is said that in macroscopic objects such as the brain, quantum fluctuations are said to be self-averaging thereby failing to contribute to the highly dynamic activity of the brain. But Jedlicka argues that there are ways by which quantum events could influence the activities of the brain thereby also increasing the brain's computational power as well as its informational coding capacities. Jedlicka considers two objections raised against the quantum brain hypothesis.

---

[30] Jedlicka, P. (2017). Revisiting the quantum brain hypothesis: Toward quantum (Neuro) biology? *Frontiers in Molecular Neuroscience, 10*(366), 1–8. For related issues, see Idris, Z., et al. (2021). Quantum and electromagnetic fields in our universe and brain: A new perspective to comprehend brain function. *Brain Sciences*, *11*(558), 1–15.

The first objection states that neuronal signaling molecules including neurons and neural networks are said to be too big for quantum phenomena to make any noticeable role in the way they function. From this, some conclude that quantum events simply average out. In this case they show their irrelevance in terms of having any meaningful impact on brain functions. The second objection states that the interactions of neuronal networks, which include their noisy, wet, and warm environment, destroy nontrivial quantum states such as entanglements.

Citing empirical evidence, Jedlicka argues that each of these objections fails to show that quantum effects play no significant role in brain function. Jedlicka argues that given that the brain is a complex nonlinear system with high sensitive to tiny quantum fluctuations, it could be the case that it can amplify small-scale quantum effects. In this regard, Jedlicka singles out two interrelated ways in which quantum events could be said to influence the activities of the brain.

First, "non-trivial quantum effects can speed up the computational processes in living organisms at the microscopic level" and second, "non-linear chaotic dynamics can amplify lowest-level quantum fluctuations upward, modulating even larger-scale mesoscopic and maybe also macroscopic neuronal activity."[31] For example, in support of the first point, Jedlicka argues that empirical evidence shows that non-trivial quantum processes have been located in living systems. In this case, long-lasting quantum coherence has been observed when it comes to electron transfer involving photosynthesis. Quantum effects are also said to be important in living things when it comes to photoreceptors, which play an important role in vision. Moreover, quantum effects are said to play a key role in the olfactory systems. Although direct experimental evidence is not available, Jedlicka argues that quantum events might affect electrical signals in the brain as well. What happens at the quantum level also has wider implications for the study of the nature of consciousness.

Second, Jedlicka argues that despite quantum effects are taken to be invisible in living organisms at the macroscopic level, small-scale effects could exert their influence on large-scale brain activity and the resulting behavior. Jedlicka further argues that the common view that says that tiny fluctuations can cancel each other out in macro-systems is not necessarily true given that the brain is a nonlinear complex system. Jedlicka claims that given the brain's complex networks of molecules, cells, microcircuits, and brain regions: small-scale fluctuations are not averaged out. By contrast, the fluctuations can be amplified. As Jedlicka puts it,

> Quantum fluctuations on the lowest level of scale may influence the initial state of the next level of scale, while the higher levels shape the boundary conditions of the lower ones. This hierarchy of nested networks with many feedback loops exploits rather than cancels out the quantum effects.[32]

---

[31] Jedlicka (2017, p. 2).

[32] Ibid., p. 4

Philosophical Issues

Jedlicka's account here can be taken as indicative of how much mere empirical studies without some sort of philosophical underpinning raise lots of questions rather than provide us with answers. Given space limitations, I will only mention two problems regarding quantum events, brain events, and mental events. I raise these issues without attempting to advance any elaborate discussion. My aim is simply to encourage interested readers to pursue solutions for them. In doing so, I also want to remind myself to revisit the issues in detail to make an informative philosophical contribution in the future.

The three sorts of events that I mentioned above require sustained philosophical machinery to show how they are related as well as how they can be said to interact with each other. However, there are two serious challenges we face in thinking about these issues that the empirical works summed up above simply fail to give us any clarity on. I shall call these problems the *distinguishability* and the *invisibility* problems, respectively.

To make sense of the *distinguishability problem*, simply think of how the brain's electrical and chemical signals work in tandem. The process seems to be straightforward when we read books about how the brain functions. However, the psycho-physiological processes themselves are no simple matter to wrap our minds around. So the sort of explanations given by Jedlicka simply does not give us any guidance as to what criterion we should use to distinguish between, say, the quantum event and the normal brain event. This raises a serious question regarding what I call ontological identity. In talking about, say, brain event and quantum event, are we using two distinct descriptions to refer to numerically the same event, or are we referring to two distinct, but uniquely identifiable events? What is the empirical evidence for this? What sort of criterion is available to undertake such a task? Would this matter be analogous to the philosophers' beloved, the *morning star* and the *evening star* scenario?

To make sense of *the invisibility problem*, remind yourself again of the detailed discussion we had undertaken concerning (Q1). If Lowe's mental causation model is on the right track, which I assume that it is, then the presence of mental events relative to brain events as well as quantum events turns out to be invisible. That means that Jedlicka's empirical-based account has very little significance if at all in giving us insights into the nature of mental events and their causal influence both on brain events and quantum events. Even if Jedlicka's account of how quantum fluctuations influence brain activity turns out to hold water, the account in question leaves out the aspects of mental events. That means for all the reasons Lowe raised via his model of mental causation, no empirical solution can be found to settle the role as well as the presence of mental events following the explanations given by Jedlicka. Moreover, the invisibility problem gives rise to another complex problem. This complex problem can be described as a *triangulation unity problem*.

What I mean by *triangulation unity problem* is this. That is, given Jedlicka's view as discussed above, we have three species of events, namely the *brain event*, the *quantum event*, and the *mental event*. How are these three species of events related? What empirical way is available for us to account for the nature of the three-way unity between these three species of events?

Q3: Would it be possible to show that the soul can affect the brain without our observing mysterious energy variations?

Engaging in any elaborate assessment of this question takes us back to what we have already discussed under (Q1). I have nothing special to say in response to this question since the question largely is a repeat of (Q1). That said, however, a reference to mysterious energy variation is something that may require a little comment.

If as we have already discussed, the unobservability of the soul is a manifestation of its very nature, then whatever its causal input turns out to be in the physical domain, it should not be cast as a mysterious input. I think referring to mysteriousness in this context simply misplaces the issue at hand. It does this by requiring the empirical observability of the soul's causal input as a precondition for us to take its operations as genuine or real.

Following Colin McGinn's style,[33] one could argue that how the soul affects the brain may never be observable. Possibly, as human beings, we are not cognitively equipped to unlock the mystery of the soul's interaction with the brain. But that does not mean that no such interaction is genuine. Rather it only means that we are not fortunate enough to unlock the mystery. But this shows only the limitations on our part as opposed to limitations on the part of the reality. In light of these and other reasons we saw under (Q1), expecting the soul's action to be less mysterious does not seem to be an effective way to think about this issue.

---

[33] For C. McGinn's well-known view of mysterianism regarding the link between mind and brain, see *The mysterious flame: Conscious minds in a material world*. (1999). NY: Basic Books.

Bibliography

Adams, R. M. (1987). *The virtues of faith*. Oxford University Press.

Addis, L. (2005). The necessity and nature of mental content. In G. Forrai & G. Kampis (Eds.), *Intentionality: Past and future* (pp. 1–14). Rodopi B.V.

Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, *33*, 147–182.

Ajana, B., Young, G., & Whitty, M. T. (2005). Disembodiment and cyberspace: A phenomenological approach. *Electronic Journal of Sociology*, *7*(2), 1–10.

Alanen, L. (2008). Descartes' mind-body composites, psychology, and naturalism. *Inquiry*, *51*(5), 464–484.

Alexander, S. (1927). *Space, time and deity* (Vol. 2). Macmillan.

Alkire, M. (2008). Consciousness and anesthesia. *Science*, *322*(5903), 876–880.

Alter, T., & Nagasawa, Y. (Eds.). (2015). *Consciousness in the physical world: Perspectives on Russellian monism*. Oxford University Press.

Anscombe, E. (1975). The first person. In S. D. Guttenplan (Ed.), *Mind and language* (pp. 45–65). Oxford University Press.

Antonietti, A. (2008). Must psychologists be dualists? In A. Antonietti, A. Corradini, & J. Lowe (Eds.), *Psycho-physical dualism today: An interdisciplinary approach* (pp. 37–68). Rowman and Littlefield.

Aquinas, T. (1259/2009). *Summa theologiæ* (Latin-English ed., Vols. *1–8*, The Fathers of the English Dominican Province, Trans.). NovaAntiqua.

Aquinas, T. (1949). *Disputed questions on spiritual creatures*. (M. C. Fitzpatrick & J. J. Wellmuth, Trans.). Marquette University Press.

Aranda-Anzaldo, A. (2016). Assuming in biology the reality of real virtuality (a come back for entelechy?). *Ludus Vitalis*, *19*(36), 333–342.

Aristotle. (1963). *Categories* (J. L. Ackrill, Trans.). Clarendon Press.

Aristotle. (1984). de Anima. In J. Barnes (Ed.), *The complete works of Aristotle* (Revised Oxford ed., Vol. 2). Princeton University Press.

Armstrong, D. M. (1968). *A materialist theory of mind*. Routledge and Kegan Paul.

Armstrong, D. M. (1978a). Naturalism: Materialism and first philosophy. *Philosophia*, *8*, 262.

Armstrong, D. M. (1978b). *A theory of universals, Vol.2 of universals and scientific realism*. Cambridge University Press.

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism,
First Edition. Brandon Rickabaugh and J. P. Moreland.

© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

Armstrong, D. M. (1989). *Universals: An opinionated introduction* (p. 55). Westview Press.

Armstrong, D. M. (1997). *A world of states of affairs*. Cambridge University Press.

Augustine. (1947a). Immortality of the soul. In L. Schopp (Ed.), *The fathers of the church* (Vol. 4). Catholic University of America Press.

Augustine. (1947b). The magnitude of the soul. In J. J. McMahon SJ. (Ed.), *The fathers of the church* (Vol. 4). Catholic University of America Press.

Augustine. (1991). *The trinity* (E. Hill & O. P. H. Park, Introduction, Translation, and Notes). New City Press.

Augustine. (1993a). *On freedom of the will* (T. Williams, Trans.). Hackett.

Augustine. (1993b). *The city of god* (M. Dods, Trans.). The Modern Library.

Augustine. (2006). *Confessions* (2nd ed., F. J. Sheed & M. P. Foley, Trans.). Hackett .

Austin, C., & Marmodoro, A. (2018). Structural powers and the homeodynamic unity of organisms. In W. M. R. Simpson, R. C. Koons, & N. J. Teh (Eds.), *Neo-Aristotelian perspectives on contemporary science*. Routledge.

Averill, E., & Keating, B. F. (1981). Does interactionism violate a law of classical physics? *Mind*, *90*(357), 102–107.

Bailey, A. (1998). The five kinds of levels of description. In S. R. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.), *Towards a science of consciousness II: The second Tucson discussions and debates*. MIT Press.

Bailey, A. (2014). You needn't be simple. *Philosophical Papers*, *43*(2), 145–160.

Bailey, A. (2016). Composition and the cases. *Inquiry*, *5*, 453–470.

Bailey, A. (2021). *Monotheism and human nature*. Cambridge University Press.

Bailey, A., Rasmussen, J., & Luke, V. H. (2011). No pairing problem. *Philosophical Studies*, *154*, 349–360.

Baker, L. R. (2000). *Persons, and bodies: A constitution view*. Cambridge University Press.

Baker, M. C., & Goetz, S. (2011). Introduction. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis: Investigation into the existence of the soul* (pp. 1–25). Continuum.

Balog, K. (2012). Acquaintance and the mind–body problem. In S. Gozzano & C. Hill (Eds.), *New perspectives on type identity: The mental and the physical* (pp. 16–42). Cambridge University Press.

Balog, K. (2019). Hard, harder, hardest. In A. Sullivan (Ed.), *Sensations, thoughts, and language: Essays in honor of Brian Loar* (pp. 265–289). Routledge.

Balz, A. G. A. (1951). *Cartesian studies*. Columbia University Press.

Bandettini, P. A. (2009). What's new in neuroimaging methods? *The Year in Cognitive Neuroscience, Annals of the New York Academy of Sciences*, *11*(56), 260–293.

Banick, K. (2021). How to be an adverbialist about phenomenal intentionality. *Synthese*, *198*, 661–686.

Barnes, G. (2002). Should property-dualists be substance-hylomorphists? *Proceedings of the American Catholic Philosophical Association*, *75*, 285–299.

Barnes, G. (2004). Is dualism religiously and morally pernicious? *American Catholic Philosophical Quarterly*, *78*(1), 99–106.

Barnett, D. (2008). The simplicity intuition and its hidden influence on philosophy of mind. *Noûs*, *42*(2), 308–335.

Barnett, D. (2010). You are simple. In R. Koons & G. Bealer (Eds.), *The waning of materialism*. Oxford University Press.

Barret, J. L. (2012). *Born believers: The science of children's religious belief*. Free Press.

Basile, P. (2009). Back to Whitehead? Galen Strawson and the rediscovery of panpsychism. In D. Skrbina (Ed.), *Mind that abides* (pp. 179–199). Benjamins.

Basile, P. (2010). It must be true—but how can it be? Some remarks on panpsychism and mental composition. *Royal Institute of Philosophy Supplement*, *67*, 93–112.

Bayne, T. (2008). The unity of consciousness and the split-brain syndrome. *The Journal of Philosophy*, *105*(6), 277–300.

Bayne, T. (2010). *The unity of consciousness*. Oxford University Press.

Bayne, T. (2018). Problems with unity of consciousness arguments for substance dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 208–225). Blackwell.

Bayne, T., & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration and dissociation* (pp. 23–58). Oxford University Press.

Bealer, G. (1993). Materialism and the logical structure of intentionality. In H. Robinson (Ed.), *Objections to physicalism* (pp. 101–126). Clarendon.

Bealer, G. (1997). Self-consciousness. *The Philosophical Review*, *106*(1), 69–117.

Bealer, G. (1998). Intuition and the autonomy of philosophy. In M. DePaul & W. Ramsey (Eds.), *Rethinking intuition. The psychology of intuition and its role in philosophical inquiry* (pp. 201–240). Rowman and Littlefield.

Bealer, G. (2002). Modal epistemology and the rationalist renaissance. In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility*. Oxford University Press.

Beauregard, M. (2007). Mind does really matter: Evidence from neuroimaging studies of emotional self-regulation, psychotherapy and placebo effect. *Progress in Neurobiology*, *81*(4), 218–236.

Bechara, A., & Tranel, D. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, *6*, 215–225.

Bechtel, W., Mandik, P., Mundale, J., Stufflebeam, R. S. (Eds.). (2001). *Philosophy and the neurosciences: A reader*. Blackwell.

Bedau, M. (2010). Weak emergence and context-sensitive reduction. In A. Corradini & T. O'Connor (Eds.), *Emergence in science and philosophy* (pp. 46–63). Routledge.

Bengson, J. (2014). How philosophers use intuition and 'intuition'. *Philosophical Studies*, *171*(3), 555–576.

Bengson, J. (2015). The intellectual given. *Mind*, *124*(495), 707–760.

Bennett, K. (2007). Mental causation. *Philosophy Compass*, *2*(2), 316–337.

Bennett, K. (2021). Why I am not a dualist. In U. Kriegel (Ed.), *Oxford studies in philosophy of mind* (Vol. 1, pp. 208–231). Oxford University Press.

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Blackwell.

Benton, M. A. (2018). God and interpersonal knowledge. *Res Philosophica*, *95*(3), 421–447.

Bergmann, G. (1967). *Realism: A critique of Brentano and Meinong*. University of Wisconsin Press.

Bergmann, G. (2004). Realism: A critique of Brentano and Meinong. In I. E. Tegtmeier (Ed.), *Collected works* (p. 3). Ontos Verlag.

Bering, J. (2002). Intuitive conceptions of dead agents' minds: The natural foundations of afterlife beliefs as phenomenological boundary. *Journal of Cognition and Culture*, *2*(4), 263–308.

Bering, J. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, *29*(5), 453–462.

Bering, J., & Bjorklund, D. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, *40*(2), 217–233.

Bering, J., Blasi, C. H., & Bjorklund, D. (2005). The development of 'afterlife' beliefs in religiously and secularly schooled children. *British Journal of Developmental Psychology*, *23*, 587–607.

Bernet, R. (2003). Desiring to know through intuition. *Husserl Studies*, *19*, 153–166.

Bickle, J. (Ed.). (2009). *The Oxford handbook of philosophy and neuroscience*. Oxford University Press.

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*, 382–386.

Bishop, J. (1989). *Natural agency*. idge University Press.

Blackmore, S., & Troscianko, E. (2018). *Consciousness: An introduction* (3rd ed.). Routledge.

Block, N. (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science*, *9*, 261–325.

Block, N. (1981). Psychologism and behaviorism. *The Philosophical Review*, *90*, 5–43.

Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, *18*, 227–247.

Block, N., & Fodor, J. (1972). What psychological states are not. *Philosophical Review*, *81*, 159–181.

Block, N., & Stalnaker, R. (1999). Conceptual analysis, dualism, and the explanatory gap. *The Philosophical Review*, *108*(1), 1–46.

Bloom, P. (2004). *Descartes's baby: How the science of child development explains what makes us human*. Basic Books.

Bloom, P. (2007). Religion is natural. *Developmental Science*, *10*, 147–151.

Bloom, P. (2009). Religious belief as an evolutionary accident. In M. Murray & J. Schloss (Eds.), *The believing primate: Scientific, philosophical, and theological reflections on the origin of religion* (pp. 118–127). Oxford University Press.

Bogardus, T. (2014). Undefeated dualism. *Philosophical Studies*, *165*, 445–466.

Bolzano, B. (1838). *Athanasia; oder Gründe für die Unsterblichkeit der Seele*. J. G. v. Seidleschen Buchhandlung.

Bostrom, N. (2008). Why I want to be a posthuman when I grow up. In B. Gordijn & R. Chadwick (Eds.), *Medical enhancement and posthumanity* (pp. 107–136). Springer.

Boyd, J. H. (1994). *Soul psychology: How to understand your soul in light of the mental health movement*. Soul Research Institute.

Braddon-Mitchel, D., & Jackson, F. (1996). *Philosophy of mind and cognition*. Blackwell.

Brandt, R., & Kim, J. (1967). The logic of the identity theory. *The Journal of Philosophy*, *64*(17), 515–537.

Bremmer, J. (1983). *The early Greek concept of the soul*. Princeton University Press.

Brentano, F. (1973). *Psychology from an empirical standpoint* (A. C. Rancurello, D. B. Terrell, & L. L. McAlister, Eds., & Trans.). Routledge.

Brentano, F. (1995). *Psychology from an empirical standpoint* (O. Kraus & L. L. McAlister, Eds.). Routledge.

Brian, P. J. (2020). Conservation laws and the philosophy of mind: Opening the black box, finding a mirror. *Philosophia*, *48*(2), 673–707.

Broad, C. D. (1925). *The mind and its place in nature*. Kegan Paul, Trench, Trubner and Co.

Broad, C. D. (1959). Autobiography. In P. A. Schlipp (Ed.), *The philosophy of C.D. Broad* (pp. 3–68). Tudor.

Broks, P. (2005). Out of mind. *Prospect*, *109*(April). https://www.prospectmagazine.co.uk/magazine/6913-outofmind

Brower, J. E. (2015). Matter. In R. Audi (Eds.), *The Cambridge dictionary of philosophy*. Cambridge University Press.

Brower, J. E. (2017). *Aquinas's ontology of the material world: Change, hylomorphism, and material objects*. Oxford University Press.

Brown, C. M. (2005). *Aquinas and the ship of Theseus*. Continuum.

Brüntrup, G. (1998). Is psycho-physical emergentism committed to dualism? The causal efficacy of emergent mental properties. *Erkenntnis*, *48*(2/3), 133–151.

Brüntrup, G., & Jaskolla, L. (Eds.). (2017). *Panpsychism: Contemporary perspectives*. Oxford University Press.

Butchvarov, P. (1995). Categories. In J. Kim & E. Sosa (Eds.), *A companion to metaphysics*. Oxford University Press.

Cain, M. J. (2016). *The philosophy of cognitive science*. Polity Press.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J., & Sirigu, A. (2004). The enjoyment of the orbitofrontal cortex in the experience of regret. *Science*, *304*, 1167–1170.

Campbell, J. (2002). *Reference and consciousness*. Clarendon.

Campbell, J. (2009). Consciousness and reference. In B. P. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford handbook of philosophy of mind*. Oxford University Press.

Campbell, K. (1984). *Body and mind* (2nd ed.). University of Notre Dame Press.

Carhart-Harris, R. L., Leec, R., Hellyer, P. J., Shanahan, M., Feilding, A., Tagliazucchi, E., Chialvo, D. R., and Nutt, D. (2014). The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in Human Neuroscience*, *8*(2), 1–22.

Chalmers, D. J. (1996). *The conscious mind*. Oxford University Press.

Chalmers, D. J. (2002). Does conceivability entail possibility? In T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility*. Oxford University Press.

Chalmers, D. J. (2010). The two-dimensional argument against materialism. In D. Chalmers, *The character of consciousness* (pp. 141–192). Oxford University Press.

Chalmers, D. J. (2017a). The combination problem for panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (pp. 179–214). Oxford University Press.

Chalmers, D. J. (2017b). Panpsychism and panprotopsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (pp. 19–47). Oxford University Press.

Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, *25*(9–10), 6–61.

Chalmers, D. J. (2020). How can we solve the meta-problem of consciousness? *Journal of Consciousness Studies*, *27*(5–6), 201–226.

Chalmers, D. J., & McQueen, K. J. (2021). *Consciousness and the collapse of the wave function*. https://arxiv.org/pdf/2105.02314.pdf

Chalmers, D. J., & McQueen, K. J. (2022). Consciousness and the collapse of the wave function. In S. Gao (Ed.), *Consciousness and quantum mechanics*. Oxford University Press.

Chisholm, R. M. (1957). *Perceiving: A philosophical study*. Cornell University Press.

Chisholm, R. M. (1969). On the observability of the self. *Philosophy and Phenomenological Research*, *30*(1), 7–21.

Chisholm, R. M. (1970). Reply to Strawson's comments. In H. E. Diefer & M. K. Munitz (Eds.), *Language, belief and metaphysics*. State University of New York Press.

Chisholm, R. M. (1977). *Theory of knowledge* (2nd ed.). Prentice-Hall.

Chisholm, R. M. (1978). Is there a mind-body problem? *Philosophical Exchange*, *9*(1), 25–34.

Chisholm, R. M. (1979). *Person and object: A metaphysics study*. Open Court.

Chisholm, R. M. (1981). *The first person: An essay on reference and intentionality*. University of Minnesota Press.

Chisholm, R. M. (1986). *Brentano and intrinsic value*. Cambridge University Press.

Chisholm, R. M. (1989a). Human freedom and the self. In R. M. Chisholm (Ed.), *On metaphysics*. University of Minnesota Press.

Chisholm, R. M. (1989b). *Theory of knowledge* (3rd ed.). Prentice-Hall.

Chisholm, R. M. (1989c). *On metaphysics*. University of Minnesota Press.

Chisholm, R. M. (1991a). On the simplicity of the soul. *Philosophical Perspectives*, *5*, 167–181.

Chisholm, R. M. (1991b). The bearers of psychological properties. *Aaiucov*, *3*, 7–14.

Chisholm, R. M. (1991c). Mind. In H. Burckhardt & B. Smith (Eds.), *Handbook of metaphysics and ontology* (pp. 544–557). Philosophia Veril.

Chisholm, R. M. (1996). *A realistic theory of categories*. Cambridge University Press.

Chudnoff, E. (2011). What intuitions are like. *Philosophy and Phenomenological Research*, *82*(3), 625–654.

Churchland, P. (1986). *Neurophilosophy: Toward a unified science of the mind/brain*. MIT Press.

Churchland, P. (2002). *Brain wise: Studies in neurophilosophy*. MIT Press.

Churchland, P. (1985). Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy*, *82*, 8–28.

Churchland, P. (1989). *A neurocomputational perspective*. MIT Press.

Churchland, P. (1995). *The engine of reason and seat of the soul*. MIT Press.

Churchland, P. (2013a). Consciousness and the introspection of 'qualitative simples'. In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience* (pp. 35–56). Springer.

Churchland, P. (2013b). *Matter and consciousness* (3rd ed.). MIT Press.

Clarke, S. (1738). Four defenses of a letter to Mr. Dodwell. In S. Clarke (Ed.), *The works of Samuel Clarke*. James, John, and Paul Knapton.

Clay, G. (2022). The contours of Locke's general substance dualism. *Journal of Modern Philosophy*, *4*(1), 9, 1–20.

Clayton, P. (2004). *Mind and emergence*. Oxford University Press.

Cleve, J. V. (1983). Conceivability and the Cartesian argument for dualism. *Pacific Philosophical Quarterly*, *64*, 35–45.

Coleman, S. (2019). Natural acquaintance. In J. Knowles & T. Raleigh (Eds.), *Acquaintance: New essays* (pp. 49–74). Oxford University Press.

Collins, R. (2011). A scientific case for the soul. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis: Investigation into the existence of the soul* (pp. 222–246). Continuum.

Conee, E., & Sider, T. (2014). *Riddles of existence: A guided tour to metaphysics*. Oxford University Press.

Conee, E., & Feldman, R. (1998). The generality problem for reliabilism. *Philosophical Studies*, *89*, 1–29.

Cooper, J. (2000). *Body, soul and life everlasting* (2nd rev. ed.). Eerdmans.

Cooper, M. J. (2007). Are we sending mixed messages? How philosophical naturalism erodes ethical instruction. *Journal of Business Ethics*, *75*, 171–180.

Copenhaver, R. (2019). Introduction to volume 4. In R. Copenhaver (Ed.), *Philosophy of mind in the early modern and modern ages: The history of the philosophy of mind* (pp. 1–15). Routledge.

Corcoran, K. (2006). *Rethinking human nature*. Baker.

Corcoran, K. (2018). Why should a Christian embrace materialism? In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 285–295). Wiley-Blackwell.

Corcoran, K., & Sharpe, K. (2016). Neuroscience and the human person. In T. Crisp & S. Porter (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology* (pp. 151–180). Eerdmans.

Corradini, A. (2008). Emergent dualism. In A. Antonietti, A. Corradini, & I. Jonathan Lowe (Eds.), *Psycho-physical dualism today: An interdisciplinary approach*. Lexington Books.

Cosmides, L., & Tooby, J. (1997). *Evolutionary psychology: A primer*. (published with University of California at Santa Barbara: Center for Evolutionary Psychology).https://scholar.google.com/scholar?q=Leda+Cosmides+and+John+Tooby,+Evolutionary+Psychology:++A+Primer&hl=en≈sdt=0≈vis=1&oi=scholarthttps://scholar.google.com/scholar?q=Leda+Cosmides+and+John+Tooby,+Evolutionary+Psychology:++A+Primer&hl=en≈sdt=0≈vis=1&oi=scholart

Crane, T. (2001). Intentional objects. *Ratio*, *14*, 336–349.

Crick, F. (1994). *The astonishing hypothesis: The scientific search for the soul*. Charles Scribner's Sons.

Cucu, A. C., & Brian Pitts, J. (2019). How dualists should (not) respond to the objection from energy conservation. *Mind and Matter*, *17*(1), 95–121.

Cummins, R., & Cummins, D. (2000). *Minds, brains, and computer: The foundations of cognitive science*. Blackwell.

Cutter, B. (forthcoming). The AI ensoulment hypothesis. *Faith and Philosophy*. Available online.

Dainton, B. (2008). *The phenomenal self*. Oxford University Press.

Daniel, D. H. (2017). Hylomorphism, new mechanisms, and explanations in biology, neuroscience, and psychology. In R. C. Koons, N. J. Teh, & W. M. R. Simpson (Eds.), *Neo-Aristotelian perspectives on contemporary science* (pp. 293–326). Routledge.

Daniel, D. H., & Meadows, G. A. (2013). Aristotle and the philosophical foundations of neuroscience. *Proceedings of the American Catholic Philosophical Association*, *87*, 213–230.

Dawkins, R. (1995). *River out of Eden: A Darwinian view of life*. Basic Books.

Demertzi, A. (2009). Dualism persists in the science of mind. In N. D. Schiff & S. Laureys (Eds.), *Disorders of consciousness: Annals of the New York Academy of Sciences, Vol. 1157* (pp. 1–9). Wiley-Blackwell.

Dempsey, L. P. (2010). An early sensation-based argument for dualism. *Locke Studies*, *10*, 159–177.

Dennett, D. C. (1984). *Elbow room*. MIT Press.

Dennett, D. C. (1987). *The intentional stance*. MIT Press.

Dennett, D. C. (1991). *Consciousness explained*. Back Bay Books.

Dennett, D. C. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. MIT Press.

Dennett, D. C. (2006). You have to give up your intuitions about consciousness. In S. Blackmore (Ed.), *Conversations about consciousness* (pp. 79–91). Oxford University Press.

Denton, M., Kumaramanickavel, G., & Legge, M. (2013). Cells as irreducible wholes: The failure of mechanism and the possibility of an organicist revival. *Biology and Philosophy*, *28*, 31–52.

Des Chene, D. (2000). *Life's form: Late Aristotelian conceptions of the soul*. Cornell University Press.

Descartes, R. (1988a). *Descartes selected philosophical writings* (Vol. 1, J. Cottingham, R. Stoothoff, & D. Murdoch, Trans.). Cambridge University Press.

Descartes, R. (1988b). *Descartes selected philosophical writings* (Vol. 2, J. Cottingham, R. Stoothoff, D. Murdoch, & A. Kenny, Trans.). Cambridge University Press.

Descartes, R. (1991). *Descartes selected philosophical writings* (Vol. 3: The Correspondence, J. Cottingham, R. Stoothoff, & D. Murdoch, Trans.). Cambridge University Press.

Dominus, S. (2011). Could conjoined twins share a mind? *New York Times*, May 25.

Double, R. (1985). Phenomenal properties. *Philosophy and Phenomenological Research*, *45*(3), 383–392.

Ducasse, C. J. (1968). Moore's refutation of idealism. In P. A. Schilpp (Ed.), *The philosophy of G.E. Moore*. Open Court.

Duncan, M. (2015). We are acquainted with ourselves. *Philosophical Studies*, *172*(9), 2531–2549.

Duncan, M. (2017). Two Russellian arguments for acquaintance. *Australasian Journal of Philosophy*, *95*(3), 461–474.

Duncan, M. (2018). Subjectivity as self-acquaintance. *Journal of Consciousness Studies*, *25*(3–4), 88–111.

Duncan, M. (2019). The self shows up in experience. *Review of Philosophy and Psychology*, *10*(2), 299–318.

Duncan, M. (2020). Knowledge of things. *Synthese*, *197*(8), 3559–3592.

Duncan, M. (2021). Acquaintance. *Philosophy Compass*, *16*(3), e12727. Online First. https://doi.org/10.1111/phc3.12727

Eberl, J. T. (2010). Varieties of dualism: Aquinas and Swinburne. *International Philosophical Quarterly*, *50*(1), 39–56.

Elliott, M. L, Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychological Science*, *31*(7), 792–806.

Ellis, B. (2002). *The philosophy of nature: A guide to the new essentialism*. Acumen Press.

Estes, D., Wellman, H. M., & Woolley, J. (1989). Children's understanding of mental phenomena. In H. Reese (Ed.), *Advances in child development and behavior* (Vol. 22). Academic Press.

Evnine, S. (2008). Modal epistemology: Our knowledge of necessity and possibility. *Philosophy Compass*, *3*(4), 665–670.

Ewing, A. C. (1934). *Idealism: A critical survey*. Methuen.

Ewing, A. C. (1973). *Value and reality*. George Allen and Unwin.

Fales, E. (2007). Naturalism and physicalism. In M. Martin (Ed.), *The Cambridge companion to atheism* (pp. 118–135). Cambridge University Press.

Feldman, R. (1995). Reliability and justification. *The Monist*, *68*(2), 159–174.

Fernandes, P., & Martin, M. (2000). *Theism vs. atheism: The internet debate*. IBD Press.

Fernandez-Duque, D., & Schwartz, B. (2016). Common sense beliefs about the central self, moral character, and the brain. *Frontiers in Psychology*, *6*(2007), 1–12.

Feser, E. (2014). *Scholastic metaphysics*. Editiones Scholasticae.

Fiala, B., Arico, A. A., & Nichols, S. (2011). On the psychological origins of dualism: Dual-process cognition and the explanation gap. In E. Slingerland & M. Collard (Eds.), *Creating consilience: Integrating the sciences and the humanities* (pp. 88–110). Oxford University Press.

Fine, K. (1995). Part-whole. In B. Smith & D. W. Smith (Eds.), *The Cambridge companion to Husserl* (pp. 463–485). Cambridge University Press.

Fodor, J. (1987). *Psychosemantics*. MIT Press.

Fodor, J. (1994). The mind-body problem. In R. Warner, T. Szubka, & U. K. Oxford (Eds.), *The mind-body problem*. Blackwell.

Forstmann, M., & Burgmer, P. (2015). Adults are intuitive mind-body dualists. *Journal of Experimental Psychology*, *144*, 222–235.

Forstmann, M., & Burgmer, P. (2017). Antecedents, manifestations, and consequences of belief in mind-body dualism. In C. Zedelius, B. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health* (pp. 181–206). Springer.

Forstmann, M., Burgmer, P., & Mussweiler, T. (2012). The mind is willing, but the flesh is weak: The effects of mind-body dualism on health behavior. *Psychological Science*, *23*, 1239–1245.

Forstmann, M., & Burgmer, P. (2018). A free will needs a free mind: Belief in substance dualism and reductive physicalism differentially predict belief in free will and determinism. *Consciousness and Cognition*, *63*, 280–293.

Foster, J. (1991). *The immaterial self: A defense of the Cartesian dualist conception of mind*. Routledge.

Fowler, D. F. (1999). *Descartes on the human soul*. Kluwer Academic.

Francescotti, R. (2001). Property dualism without substance dualism? *Philosophical Papers*, *30*(2), 93–116.

Francken, J. C., Beerendonk, L., Molenaar, D., Fahrenfort, J. J., Kiverstein, J. D., Seth, A. K., & van Gaal, S. (2021). *An academic survey on theoretical foundations, common assumptions and the current state of the field of consciousness science*. *2021*, PsyArXiv *June* 14. 1–13.

Frege, G. (1965). The thought: A logical inquiry. *Mind*, *65*(295), 289–311.

Friberg-Fernros, H. (2020). The explanatory power of the soul: Why liberals should seriously consider dualism. *Journal of Religious Ethics*, *48*(1), 101–121.

Fumerton, R. (2013). *Knowledge, thought, and the case for dualism*. Cambridge University Press.

Fumerton, R. (2019). Acquaintance: The foundation of knowledge and thought. In J. Knowles & T. Raleigh (Eds.), *Acquaintance: New essays*. Oxford University Press.

Ganssle, G. E. (2022). Divine causation and the pairing problem. In G. E. Ganssle (Ed.), *Divine causation*. Routledge.

Garcia, R. K. (2014a). Descartes's independence conception of substance and his separability argument for substance dualism. *Journal of Philosophical Research*, *39*, 165–190.

Garcia, R. K. (2014b). Closing in on causal closure. *Journal of Consciousness Studies*, *21*(1–2), 96–109.

Gasser, G., & Stefan, M. (2012). *Personal identity: Complex or simple?* Cambridge University Press.

Gaymer, M. G. (2017). *Quantum physics*. Oxford University Press.

Gendler, T. S., & Hawthorne, J. (Eds.). (2002). *Conceivability and possibility*. Clarendon.

Gertler, B. (2011). *Self-Knowledge*. Routledge.

Gertler, B. (2012a). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 89–123). Oxford University Press.

Gertler, B. (2012b). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 93–127). Oxford University.

Gertler, B. (2020). Dualism: How epistemic issues drive debates about the ontology of consciousness. In U. Kriegel (Ed.), *The Oxford handbook of the philosophy of consciousness* (pp. 277–300). Oxford University Press.

Ginet, C. (1990). *On action*. University Press.

Goetz, S. (2001). Modal dualism: A critique. In K. Corcoran (Ed.), *Soul, body and survival*. Cornell University Press.

Goetz, S., & Taliaferro, C. (2011). *A brief history of the soul*. Wiley-Blackwell.

Goff, P. (2015). Real acquaintance and physicalism. In P. Coates & S. Coleman (Eds.), *Phenomenal qualities: Sense, perception, and consciousness* (pp. 121–146). Oxford University Press.

Goff, P. (2017). *Consciousness and fundamental reality*. Oxford University Press.

Goff, P. (2019). Did the universe design itself? *International Journal for Philosophy of Religion*, *85*(1), 99–122.

Goldman, A. (1979). What is justified belief? In G. Pappas (Ed.), *Justification and knowledge* (pp. 1–25). D. Reidel.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. MIT Press.

Gottlieb, S., & Lombrozo, T. (2018). Can science explain the human mind? Intuitive judgments about the limits of science. *Psychological Science*, *29*(1), 121–130.

Grahek, N. (2007). *Feeling pain and being in pain* (2nd ed.). MIT Press.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400.

Grossmann, R. (1983). *The categorial structure of the world*. Indiana University Press.

Grossmann, R. (1992). *The existence of the world*. Routledge.

Guta, M. P. (2011a). Frank Jackson's location problem and argument from the self. *Philosophia Christi*, *1*(13), 35–58.

Guta, M. P. (2011b). Neuroscience or neuroscientism? *Perspectives on Science and Christian Faith*, *63*(1), 69–70.

Guta, M. P. (2015). Consciousness, first-person perspective, and neuroimaging. *Journal of Consciousness Studies*, *22*(11–12), 218–245.

Guta, M. P. (Ed.). (2019). *Consciousness and the ontology of properties*. Routledge.

Guta, M. P. (2021). Metaphysics, natural science and theological claims: E. J. Lowe's approach. *TheoLogica: An International Journal for Philosophy of Religion and Philosophical Theology*, *5*(2), 129–160.

Guta, M. P. (work in preparation). *Mirror neurons, consciousness and the bearer question*.

Guta, M. P. (work in preparation). *In what sense is consciousness a property?*

Haan, E. H. F. (2020). Unified tactile detection and localization in split-brain patients. *Cortex*, *124*, 217–223.

Haldane, J. J. (1989). Naturalism and the problem of intentionality. *Inquiry*, *32*(3), 305–322.

Haldane, J. J. (2003). (I Am) thinking. *Ratio XVI*, *2*(June), 124–139.

Hall, T. W. (1998). The soul or substantive self as experiencer, actualizer, and representative in psychoanalytic theory. *Journal of Psychology and Theology*, *26*(1), 55–65.

Hardcastle, V. G., & Stewart, C. M. (2009). fMRI: A modern cerebrascope? The case of pain. In J. Pickle (Ed.), *The Oxford handbook of philosophy and neuroscience*. Oxford University Press.

Harrison, G. K. (2016). A moral argument for substance dualism. *Journal of the American Philosophical Association*, *1*, 21–35.

Hart, W. D. (1988). *The engine of the soul*. Oxford University Press.

Hart, W. D. (1994). Dualism. In S. Guttenplan (Ed.), *A companion to the philosophy of mind* (pp. 265–270). Blackwell.

Hasker, W. (1999). *The emergent self*. Cornell University Press.

Hasker, W. (2005). On behalf of emergent dualism. In J. B. Green & S. L. Palmer (Eds.), *In search of the soul: Four views on the mind-body problem* (pp. 75–100). InterVarsity Press.

Hasker, W. (2010). Persons and the unity of consciousness. In R. Koons & G. Bealer (Eds.), *The waning of materialism* (pp. 175–190). Oxford University Press.

Hasker, W. (2012). Is materialism equivalent to dualism? In B. P. Göcke (Ed.), *After physicalism* (pp. 180–199). University of Notre Dame Press.

Hasker, W. (2013). The dialectic of body and soul. *American Catholic Philosophical Quarterly*, *87*(3), 495–509.

Hasker, W. (2016). Do my quarks enjoy Beethoven? In T. M. Crisp, S. L. Porter, & T. E. Gregg (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology*. William B. Eerdmans.

Hasker, W. (2018). An emergent dualism response. In J. J. Loose, J. L. Angus, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism*. Wiley-Blackwell.

Hawthorne, J., & McDonough, J. K. (1998). Numbers, minds, and bodies: A fresh look at mind-body dualism. *Philosophical Perspectives*, *12*, 349–371.

Heil, J. (2012). *Philosophy of mind: A contemporary introduction* (3rd ed.). Routledge.

Henrich, D. (1982). Fichte's original insight. Edited by D.E. Christensen. *Contemporary German Philosophy*, *1*, 15–53.

Henry, J. (2004). Metaphysics and the origins of modern science: Descartes and the importance of laws of nature. *Early Science and Medicine*, *9*(2), 73–114.

Herodotus. *Historiae.*, with an English translation by A. D. Godley. Cambridge. Harvard University Press. 1920. Written c. 430 BC.

Hoffman, D. D. (2008). Conscious realism and the mind-body problem. *Mind and Matter*, *6*(1), 87–121.

Hoffman, J., & Rosenkrantz, G. (1991). Are souls unintelligible? *Philosophical Perspectives*, *5*, 183–212.

Hoffman, J., & Rosenkrantz, G. S. (1994). *Substance among other categories*. Cambridge University Press.

Hoffman, J., & Rosenkrantz, G. S. (1997). *Substance: Its nature and existence*. Routledge.

Hoffman, J., & Rosenkrantz, G. S. (2002). *The divine attributes*. Blackwell.

Hoffman, P. (1986). The unity of Descartes' man. *The Philosophical Review*, *95*, 339–369.

Holstrom, N. (1977). Firming up soft determinism. *The Personalist*, *58*, 39–51.

Hopkins, P. (2008). A moral vision for transhumanism. *Journal of Evolution and Technology*, *19*(1), 3–7.

Hopp, W. (2011). *Perception and knowledge*. Cambridge University Press.

Hopp, W. (2015). Phenomenal conservatism and the principle of all principles. In D. O. Dahlstrom, A. Elpidorou, & W. Hopp (Eds.), *Philosophy of mind and phenomenology: Conceptual and empirical approaches* (pp. 180–202). Routledge.

Hopp, W. (2020). *Phenomenology: An introduction*. Routledge.

Horgan, T. (1993). Nonreductive materialism and the explanatory autonomy of psychology. In S. Wagner & R. Warner (Eds.), *Naturalism: A critical appraisal* (pp. 309–330). University of Notre Dame Press.

Horst, S. (2007). *Beyond reduction: Philosophy of mind and post-reductionist philosophy of science*. Oxford University Press.

Horst, S. (2009). Naturalisms in philosophy of mind. *Philosophy Compass*, *4*(1), 219–254.

Hudson, H. (2001). *A materialist metaphysics of the human person*. Cornell University Press.

Huemer, M. (2009). When is parsimony a virtue? *Philosophical Quarterly*, *59*, 216–235.

Huemer, M. (2018). *Paradox lost: Logical solutions to ten puzzles of philosophy*. Palgrave Macmillan.

Hull, D. (1989). *The metaphysics of evolution*. State University of New York Press.

Humanity+ (2020). *Transhumanism FAQ 3.0*. Online Document. https://www.humanityplus.org/transhumanist-faq

Hume, D. (1964). *A treatise of human nature*. Dent.

Hume, D. (2000). *A treatise of human nature*. (D. F. Norton & M. J. Norton, Eds.). Oxford University Press.

Husserl, E. (1900/1970). *Logical investigations* (Translation of 2nd ed., J. N. Findlay). Routledge and Kegan Paul.

Husserl, E. (1982). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy. First book: General introduction to a pure phenomenology* (F. Kersten, Trans.). Nijhoff.

Husserl, E. (1983). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy—First book: General introduction to a pure phenomenology*. (F. Kersten, Trans.). Kluwer Academic.

Husserl, E. (1989). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy—Second book: Studies in the phenomenology of constitution* (R. Rojcewicz & A. Schuwer, Trans.). Kluwer Academic.

Hutteman, A. (2004). *What's wrong with microphysicalism*. Routledge.

Idris, Z., Zakaria, Z., Yee, A. S., Fitzrol, D. N., Ghani, A. R. I., Abdullah, J. M., Wan Hassan, W. M. N., Manaf, A.A., & Chong Heng, R.O. (2021). Quantum and electromagnetic fields in our universe and brain: A new perspective to comprehend brain function. *Brain Sciences*, *11*(558), 1–15.

Inagaki, K., & Hatano, G. (1991). Constrained person analogy in young children's biological inference. *Cognitive Development*, *6*, 219–231.

Inman, R. D. (2018). *Substance and the fundamentality of the familiar: A neo-Aristotelian mereology*. Routledge.

Jackson, F. (1982a). Epiphenomenal qualia. *Philosophical Quarterly*, *32*, 127–136.

Jackson, F. (1982b). What Mary didn't know. *Journal of Philosophy*, *83*, 291–295.

Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy*, *83*, 291–295.

Jackson, F. (1998). *From metaphysics to ethics*. Oxford University Press.

James, S. (2000). The emergence of the Cartesian mind. In T. Crane & S. Paterson (Eds.), *History of the mind body problem*. Routledge.

James, V. C. (1983). Conceivability and the Cartesian argument for dualism. *Pacific Philosophical Quarterly*, *64*, 35–45.

James, W. (1890). *The principles of psychology*. H. Holt.

James, W. (1904). Does 'consciousness' exist? *The Journal of Philosophy, Psychology and Scientific Methods*, *1*(18), 477–49.

Jansson, L., & Tallant, J. (2017). Quantitative parsimony: Probably for the better. *British Journal for the Philosophy of Science*, *68*, 781–803.

Jedlicka, P. (2017). Revisiting the quantum brain hypothesis: Toward quantum (neuro) biology? *Frontiers in Molecular Neuroscience*, *10*(366), 1–8.

Johnson, S. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences*, *4*, 22–28.

Jubien, M. (1997). *Contemporary metaphysics*. Blackwell.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, *93*, 1449–1475.

Kane, R. (1996). *The significance of free will*. University Press.

Kane, R. (2005). *A contemporary introduction to free will*. Oxford University Press.

Kaplan, D. (1970). On the logic of demonstratives. *Journal of Philosophical Logic*, *8*, 81–98.

Kasmier, D. (2003). Husserl's theory of a priori knowledge: A response to the failure of contemporary rationalism. Ph.D. Diss., University of Southern California.

Keas, M. (2018). Systematizing the theoretical virtues. *Synthese*, *195*, 2761–2793.

Kelly, T. (2008). Common sense as evidence: Against revisionary ontology and skepticism. *Midwest Studies in Philosophy*, *32*, 53–78.

Kim, J. (1998). *Mind in a physical world*. MIT Press.

Kim, J. (2001). Lonely souls: Causality and substance dualism. In K. Corcoran (Ed.), *Soul, body and survival: Essays on the metaphysics of human persons* (Vol. 30). Cornell University Press.

Kim, J. (2005). *Physicalism, or something near enough*. Princeton University Press.

Kim, J. (2006). *Philosophy of mind* (2nd ed.). Westview Press.

Kim, J. (2011). *Philosophy of mind* (3rd ed.). Westview Press.

Kim, J. (2018a). Against Cartesian dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 152–167). Wiley-Blackwell.

Kim, J. (2018b). Against Cartesian dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 154–167). Wiley-Blackwell.

Kim, J. (1973). Causation, nomic subsumption, and the concept of event. *Journal of Philosophy*, *70*, 217–236.

Kind, A. (2019). The mind-body problem in 20th century philosophy. In A. Kind (Ed.), *Philosophy of mind in the twentieth and twenty-first centuries: A history of the philosophy of mind* (pp. 6–73). Routledge.

Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and Cognitive Sciences*, *7*, 67–83.

Koivuniemi, M., & Curley, E. (2015). Descartes on the mind–body union: A different kind of dualism. In D. Garber & D. Rutherford (Eds.), *Oxford studies in early modern philosophy* (Vol. VII, pp. 83–112). Oxford University Press.

Koons, R. C. (2014). Staunch vs. fainthearted hylomorphism: Towards an Aristotelian account of composition. *Res Philosophica*, *91*(2), 151–177.

Koons, R. C. (2018). Against emergent individualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 377–393). Wiley-Blackwell.

Koons, R. C. (2022). Essential thermochemical and biological powers. In W. M. R. Simpson, R. C. Koons, & J. Orr (Eds.), *Neo-Aristotelian metaphysics and the theology of nature* (pp. 66–93). Routledge.

Koslicki, K. (2008). *The structure of objects*. Oxford University Press.

Kriegel, U. (2007). Intentional inexistence and phenomenal intentionality. *Philosophical Perspectives*, *21*(1), 304–307.

Kriegel, U. (2008). The dispensability of (Merely) intentional objects. *Philosophical Studies*, *141*, 79–95.

Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. Oxford University Press.

Kuhlmeier, V. A., Bloom, P., & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition*, *94*(1), 95–103.

Kuhn, T. (1977). *The essential tension: Selected studies in scientific tradition and change*. The University of Chicago Press.

Lamme, V. A. F. (2006). Towards a true neutral stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494–501.

Langsam, H. (2001). Strategy for dualists. *Metaphilosophy*, *32*(4), 395–418.

Langsam, H. (2002). Consciousness, experience, and justification. *Canadian Journal of Philosophy*, *32*(1), 1–28.

Lapointe, S. (2018). Bolzano's philosophy of mind and action. In S. Lapointe (Ed.), *Philosophy of mind in the nineteenth century: The history of philosophy of mind* (Vol. 5, pp. 42–59). Routledge.

Larmer, R. (1986). Mind-body interaction and the conservation of energy. *International Philosophical Quarterly*, *26*(3), 277–285.

Larmer, R. (2008). Miracles, physicalism, and the laws of nature. *Religious Studies*, *44*(2), 149–159.

Larmer, R. (2009). Divine agency and the principle of the conservation of energy. *Zygon*, *44*(3), 543–557.

LaRock, E. (2008). Is consciousness really a brain process? *International Philosophical Quarterly*, *48*(2), 208.

LaRock, E. (2016). Neuroscience and the hard problem of consciousness. In T. M. Crisp, S. Porter, & G. A. Ten Elshof (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology* (pp. 151–180). Eerdmans.

LaRock, E., & Collins, R. (2008). Saving our souls from materialism. *Neuroscience and the Soul*, 137–146; Eric LaRock, Is consciousness really a brain process?. *International Philosophical Quarterly*, *48*(2), 201–222.

Lavazza, A., & Robinson, H. (2016). Introduction. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 1–14). Routledge.

Lebowitz, M. S. (2014). Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice*, *21*, 67–83.

Lecon, M. (2013). Francis Suarez on the efficiency of substantial forms. *The Review of Metaphysics*, *67*(September), 107–124.

Legrenzi, P., & Umiltà, C. (2011). *Neuromania: On the limits of brain science* (F. Anderson, Trans.). Oxford University Press.

Leibniz, G. W. (1902). Leibniz to Arnauld, October 6, 1687. In G. Montgomery (Ed.), *G. W. Leibniz, discourse on metaphysics, correspondence with Arnauld, and monadology*. Open Court.

Leibniz, G. W. (1989). Leibniz, philosophical essays. In R. Ariew & D. Garber (Eds.), *Principles nature*. Hackett.

Leibniz, G. W. (1997). New system of the nature of substances and their communication, and the union which exists between the soul and the body. In R. S. Woolhouse & R. Francks (Eds.), *Leibniz's 'new system' and associated contemporary texts* (pp. 7–36). Oxford University Press.

Lennon, T., & Stainton, R. (Eds.). (2008). *The Achilles of rationalist psychology*. Springer.

Leotti, L. A., Iyengar, S. S., & Ochsner, K. N. (2010). Born to choose: The origins and value of the need for control. *Trends in Cognitive Sciences*, *14*, 457–463.

Levin, S. B., & Bliss, P. (2021). *The failed promise of transhumanism*. Oxford University Press.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, *64*, 354–361.

Levine, J. (2019). Acquaintance is consciousness and consciousness is acquaintance. In J. Knowles & T. Raleigh (Eds.), *Acquaintance: New essays* (pp. 33–48). Oxford University Press.

Levitin, D. J. (2006). *This is your brain on music: The science of human obsession*. Dutton.

Levy, N. (2016). Naturalism and free will. In K. J. Clark (Ed.), *The Blackwell companion to naturalism* (pp. 305–318). John Wiley.

Lewis, C. I. (1946). *An analysis of knowledge and valuation*. The Open Court.

Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, *74*(4), 549–567.

Lewontin, R. C. (1992). The dream of the human genome. *New York Review of Books*, May 28. https://www.nybooks.com/articles/1992/05/28/the-dream-of-the-human-genome

Lim, D. (2015). *God and mental causation*. Springer.

Loar, B. (1997). Phenomenal states. In N. Block, O. Flanagan, & G. Guzeldere (Eds.), *The nature of consciousness* (pp. 597–616). MIT Press.

Loar, B. (2004). Phenomenal states (Revised Version). In P. Ludlow, Y. Nagasawa, & D. Stoljar (Eds.), *There's something about Mary: Essays on phenomenal consciousness and Frank Jackson's knowledge argument* (pp. 219–240). MIT Press.

Lotze, H. (1887). *Metaphysics in three books: Ontology, cosmology, and psychology* (B. Bosanquet, Trans. & Ed., 2nd ed., Vol. 2). Clarendon Press.

Lotze, H. (1888). *Microcosmus: An essay concerning man and his relation to the world* (E. Hamilton & E. E. C. Jones, Trans. from the German, 3rd ed., Vol. 1). T&T Clark.

Loux, M. J., & Crisp, T. M. (2017). *Metaphysics: A contemporary introduction*. Routledge.

Lovejoy, A. O. (1936). *The great chain of being: A study of the history of an idea*. Harvard University Press.

Lovejoy, A. O. (1930). *The revolt against dualism: An inquiry concerning the existence of ideas*. Open Court.

Lowe, E. J. (1995). There are no easy problems of consciousness. *Journal of Consciousness Studies*, *2*(3), 206–271.

Lowe, E. J. (1996). *Subjects of experiences*. Cambridge University Press.

Lowe, E. J. (2000). *An introduction to the philosophy of mind*. Cambridge University Press.

Lowe, E. J. (2002). *Survey of metaphysics*. Oxford University Press.

Lowe, E. J. (2006a). *The four-category ontology*. Oxford University Press.

Lowe, E. J. (2006b). Non-cartesian dualism and the problem of mental causation. *Erkenntnis*, *65*, 5–23.

Lowe, E. J. (2008). *Personal agency: The metaphysics of mind and action*. Oxford University Press.

Lowe, E. J. (2012). The probable simplicity of personal identity. In G. Gasser & M. Stefan (Eds.), *Personal identity: Complex or simple?* Cambridge University Press.

Lowe, E. J. (2018). Non-Cartesian substance dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 168–182). Wiley-Blackwell.

Lund, D. H. (1990). Disembodied existence, personal identity, and the first-person perspective. *Idealistic Studies*, *20*(3), 187–202.

Lund, D. H. (2005). *The conscious self: The immaterial center of subjective states*. Humanity Books.

Lund, D. H. (2014). Materialism, dualism, and the conscious self. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense*. Routledge.

Lycan, W. G. (2002). Explanation and epistemology. In P. Moser (Ed.), *The Oxford handbook of epistemology* (pp. 408–433). Oxford University Press.

Lycan, W. G. (2009). Giving dualism its due. *Australasian Journal of Philosophy*, *87*(4), 551–556.

Lycan, W. G. (2013). Is property dualism better off than substance dualism? *Philosophical Studies*, *164*(2), 533–542.

Lycan, W. G. (2018). Redressing substance dualism. In J. Loose, A. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 22–40). Wiley-Blackwell.

Lyons, W. G. (2001). *Matters of the mind*. Routledge.

Madden, R. (2015). The naive topology of the conscious subject. *Noûs*, *49*(1), 55–70.

Madell, G. (1981). *The identity of the self*. The University Press.

Madell, G. (2010). The road to substance dualism. *Royal Institute of Philosophy Supplement*, *67*, 45–60.

Madell, G. (2015). *The essence of the self: In defense of the simple view of personal identity*. Routledge.

Manzotti, R., & Moderato, P. (2014). Neuroscience: Dualism in disguise. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 81–97). Routledge.

Markosian, N. (1998). Simples. *Australasian Journal of Philosophy*, *76*, 332–340.

Marmodoro, A. (2021). Hylomorphic unity. In R. Bliss & J. T. M. Miller (Eds.), *The Routledge handbook of metametaphysics* (pp. 284–299).

Marmodoro, A., & Cartwright, S. (2018). Introduction. In A. Marmodoro & S. Cartwright (Eds.), *A history of mind and body in late antiquity* (pp. 1–12). Cambridge University Press.

Marmodoro, A., & Cartwright, S. (Eds.). (2018). *A history of mind and body in late antiquity*. Cambridge University Press.

Martin, C. B. (2007). *The mind in nature*. Clarendon Press.

Martin, R., & Barresi, J. (2006). *The rise and fall of soul and self: An intellectual history of personal identity*. Columbia University Press.

Matheson, C. (2009). Review of *The mind in nature*. *History of Philosophy and Logical Analysis*, *12*(1), 315–320.

Maudlin, T. (2019). *Philosophy of physics: Quantum theory*. Princeton University Press.

Maudlin, T., Okon, E., & Sudarsky, D. (2020). On the status of conservation laws in physics: Implications for semiclassical gravity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *69*, 67–81. https://doi.org/10.1016/j.shpsb.2019.10.004

Mayr, E. (1970). *Populations, species, and evolution*. Harvard University Press.

McCarthy, J. C. (1979). Ascribing mental qualities to machines. In M. Ringle (Ed.), *Philosophical perspectives in artificial intelligence*. Humanities Press.

McCarthy, J. C. (1992). Parts' whole, and the forms of life: Husserl and the new biology. In L. Hardy & L. Embree (Eds.), *Phenomenology of natural science* (pp. 135–156). Kluwer Academic.

McGinn, C. (1989a). Can we solve the mind-body problem? *Mind*, *98*, 349–366.

McGinn, C. (1989b). *Mental content*. Basil Blackwell.

McGinn, C. (1999). *The mysterious flame: Conscious minds in a material world*. Basic Books.

Meixner, U. (2004). *The two sides of being: A reassessment of psycho-physical dualism*. Mentis-Verlag.

Meixner, U. (2006). Classical intentionality. *Erkenntnis*, *65*, 25–45.

Meixner, U. (2010). Materialism does not save the phenomena—and the alternative which does. In R. C. Koons & G. Bealer (Eds.), *The waning of materialism* (pp. 417–437). Oxford University Press.

Meixner, U. (2017). Substance dualism. In Y. Nagasawa & B. Matheson (Eds.), *The Palgrave handbook of the afterlife* (pp. 277–295). Palgrave Macmillan.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 83–99.

Menuge, A. (2009). Is downward causation possible? How the mind can make a physical difference. *Philosophia Christi*, *11*(1), 93–110.

Menuge, A., & Ganssle, G. E. (Eds.). (2022). *Philosophical essays on divine causation*. Routledge.

Merricks, T. (1994). A new objection to *a priori* arguments for dualism. *American Philosophical Quarterly*, *31*(1), 81–85.

Merricks, T. (2001). How to live forever without saving your soul. In K. Corcoran (Ed.), *Soul, body and survival*. Cornell University Press.

Michel, A. (2015). Dualism at work: The social circulation of embodiment theories in use. *Signs and Society*, *3*(S1), 41–69.

Mijuskovic, B. L. (1974). *The Achilles of rationalist arguments: The simplicity, unity, and identity of thought and soul from the Cambridge Platonists to Kant: A study in the history of an argument*. Martinus Nijhoff.

Millière, R. (2017). Looking for the self: Phenomenology, neurophysiology and philosophical significance of drug-induced ego dissolution. *Frontiers in Human Neuroscience*, *11*(245), 1–22.

Miresco, M., & Kirmayer, L. (2006). The persistence of mind-brain dualism in psychiatric reasoning about clinical scenarios. *American Journal of Psychiatry*, *163*(5), 913–918.

Monteverde, S., & Rid, A. (2012). Controversies in the determination of death: Perspectives from Switzerland. *Swiss Medical Weekly*, *142*, w13667.

Monton, B. (2018). *The argument from simplicity and the argument from induction: Atheistic induction by Boltzmann Brains*. (J. L. Walls & T. Dougherty, Eds.). Oxford University Press.

Moreland, J. P. (1998). Restoring the substance of the soul to psychology. *Journal of Psychology and Theology*, *26*(1), 29–43.

Moreland, J. P. (1999). Madell's rejection of a substantial, immaterial self. *Philosophia Christi*, *1*(1), 111–114.

Moreland, J. P. (2000). Naturalism and the ontological status of properties. In W. L. Craig & J. P. Moreland (Eds.), *Naturalism: A critical analysis* (pp. 67–109). Routledge.

Moreland, J. P. (2001). *Universals*. McGill-Queen's.

Moreland, J. P. (2002). The absurdities of Mormon materialism: A reply to the neglected Orson Pratt. In F. Beckwith, C. Mosser, & P. Owen (Eds.), *The new Mormon challenge* (pp. 243–270). Zondervan.

Moreland, J. P. (2002). Naturalism, nominalism, and Husserlian moments. *The Modern Schoolman*, *79*(January/March), 199–216.

Moreland, J. P. (2003). The knowledge argument revisited. *International Philosophical Quarterly*, *2*(170), 219–238.

Moreland, J. P. (2008). *Consciousness and the existence of God*. Routledge.

Moreland, J. P. (2009). *The recalcitrant Imago Dei: Human persons and the failure of naturalism*. SCM Press.

Moreland, J. P. (2011a). Oppy on the argument from consciousness: A rejoinder. *European Journal for Philosophy of Religion*, *3*(1), 213–226.

Moreland, J. P. (2011b). Substance dualism and the argument from self-awareness. *Philosophia Christi*, *13*(1), 21–34.

Moreland, J. P. (2012). *Christianity, neuroscience, and dualism*. (J. B. Stump & A. G. Padgett, Eds.). Blackwell.

Moreland, J. P. (2013a). A conceptualist argument for substance dualism. *Religious Studies*, *49*, 35–43.

Moreland, J. P. (2013b). Exemplification and constituent realism: A clarification and modest defense. *Axiomathes*, *23*, 247–259.

Moreland, J. P. (2014). *The soul: How we know it's real and why it matters*. Moody.

Moreland, J. P. (2016). Why top-down causation does not supply adequate support for mental causation. In T. Crisp & S. Porter (Eds.), *Neuroscience and the soul: The human person in philosophy, science, and theology* (pp. 51–73). Eerdmans.

Moreland, J. P. (2018a). In defense of a Thomistic-like dualism. In J. Loose, A. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to mind-body dualism* (pp. 102–122). Blackwell.

Moreland, J. P. (2018b). The fundamental limitations of cognitive neuroscience for stating and solving the ubiquitous metaphysical issues in philosophy of mind. *Philosophia Christi*, *20*(1), 43–51.

Moreland, J. P. (2018c). Substance dualism and the diachronic/synchronic unity of consciousness. In R. K. Loftin & J. R. Farris (Eds.), *Christian physicalism? Philosophical theological criticisms* (pp. 43–74). Lexington Books.

Moreland, J. P. (2020). Wielenberg and emergence: Borrowed capital on the cheap. In A. L. Johnson (Ed.), *God and morality: What is the best account of objective moral values and duties?* (pp. 93–114). Routledge.

Moreland, J. P. (2021). *A simple guide to experience miracles*. Zondervan.

Moreland, J. P., & Rae, S. B. (2000). *Body and soul: Human nature and the crisis in ethics*. InterVarsity Press.

Mulligan, K., & Smith, B. (1985). Franz Brentano on the ontology of mind. *Philosophy and Phenomenological Research*, *45*(4), 627–644.

Mumford, S. (2013). The power of power. In R. Groff & J. Greco (Eds.), *Powers and capacities in philosophy: The new Aristotelianism* (pp. 9–24). Routledge.

Murphy, N. (1998). Human nature: Historical, scientific and religious issues. In W. S. Brown, N. Murphy, & H. Newton Malony (Eds.), *Whatever happened to the soul?* (pp. 1–30). Fortress Press.

Murphy, N. (2005). Nonreductive physicalism. In J. B. Green & S. L. Palmer (Eds.), *In search of the soul*. InterVarsity Press.

Murphy, N. (2007). Reductionism: How did we fall into it and can we emerge from it? In N. Murphy & W. R. Stoeger (Eds.), *Evolution and emergence*. Oxford University Press.

Murphy, N. (2018). For nonreductive physicalism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism*. Wiley-Blackwell.

Murphy, N. (2006). *Bodies and souls, or spirited bodies?* Cambridge University Press.

Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., Ross, L. T. (2014). The free will inventory: Measuring beliefs about agency and responsibility. *Consciousness and Cognition*, *25*, 27–41.

Nadelhoffer, T., Murray, S., & Murry, E. (forthcoming). Intuitions about free will and the failure to comprehend determinism. *Erkenntnis*. https://doi.org/10.1007/s10670-021-00465-y

Nadelhoffer, T., & Tocchetto, D. G. (2013). The potential dark side of believing in free will (and related concepts). In G. Caruso (Ed.), *Exploring the illusion of free will and moral responsibility* (pp. 121–141). Lexington Books.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450.

Nagel, T. (1997). *The last word*. Oxford University Press.

Nagel, T. (2012). *Mind and cosmos*. Oxford University Press.

Nathan, N. M. L. (2011). Substance dualism fortified. *Philosophy*, *86*(2), 201–211.

Neuhouser, F. (1990). *Fichte's theory of subjectivity*. Cambridge University Press.

Nida-Rümelin, M. (2007). Grasping phenomenal properties. In T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge* (pp. 307–336). Oxford University Press.

Nida-Rümelin, M. (2013). The argument for subject body dualism from transtemporal identity defended. *Philosophy and Phenomenological Research*, *86*(3), 702–714.

Nida-Rümelin, M. (2017). Self-awareness. *Review of Philosophy and Psychology*, *8*, 55–82.

Nida-Rümelin, M. (2018). The experience property framework: A misleading paradigm. *Synthese*, *195*, 3361–3387.

Nolan, D. (1997). Quantitative parsimony. *British Journal for the Philosophy of Science*, *48*, 329–343.

Noonan, H. (2012). Personal identity and perplexities. In G. Gasser & M. Stefan (Eds.), *Personal identity: Complex or simple?* Cambridge University Press.

Nordan, J. (2007). *Understanding the brain*. The Teaching Company. Lecture Series, Parts 1–3.

O'Conaill, D. (2022). *Substance (elements in metaphysics)*. Cambridge University Press.

O'Connor, T. (2000). *Persons and causes*. Oxford University Press.

O'Connor, T. (2001). Dualist and agent-causal theories. In R. Kane (Ed.), *The Oxford handbook on free will* (pp. 337–355). Oxford University Press.

O'Connor, T. (2012). *Theism and ultimate explanation*. Wiley-Blackwell.

O'Connor, T. (2021). Emergent properties. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter ed.). https://plato.stanford.edu/archives/win2021/entries/properties-emergent

O'Connor, T., & Kimble, K. (2011). The argument from consciousness revisited. In J. Kvanvig (Ed.), *Oxford studies in the philosophy of religion* (Vol. 3, pp. 110–141). Oxford University Press.

O'Connor, T., & Wong, H. Y. (2005). The metaphysics of emergence. *Noûs*, *39*(4), 658–678.

Oderberg, D. S. (2007). *Real essentialism*. Routledge.

Oderberg, D. S. (2011). Essence and properties. *Erkenntnis*, *75*, 85–111.

Oderberg, D. S. (2018). The great unifier: Form and the unity of the organism. In W. M. R. Simpson, R. C. Koons, & N. J. Teh (Eds.), *Neo-Aristotelian perspectives on contemporary science*. Routledge.

Oderberg, D. S. (2022). Restoring the hierarchy of being. In W. M. R. Simpson, R. C. Koons, & J. Orr (Eds.), *Neo-Aristotelian metaphysics and the theology of nature* (pp. 94–124). Routledge.

Olson, E. T. (1997). *The human animal: Personal identity without psychology*. Oxford University Press.

Olson, E. T. (2007a). There is no problem of the self. In B. Gertler & L. Shapiro (Eds.), *Arguing about the mind* (pp. 262–277). Routledge.

Olson, E. T. (2007b). *What are we? A study in personal ontology*. Oxford University Press.

Olson, E. T. (2019). The metaphysics of artificial intelligence. In M. P. Guta (Ed.), *Consciousness and the ontology of properties* (pp. 67–84). Routledge.

Oppy, G. (2011). Review of J.P. Moreland's the argument from consciousness: A theistic argument. *European Journal for Philosophy of Religion*, *3*, 193–212.

Oppy, G. (2013). *The best argument against God*. Palgrave Macmillan.

Owen, M. (2021). *Measuring the immeasurable mind: Where contemporary neuroscience meets the Aristotelian tradition*. Lexington Books.

Page, B. (2020). Arguing to theism from consciousness. *Faith and Philosophy*, *37*(3), 336–362.

Papineau, D. (1993). *Philosophical naturalism*. Blackwell.

Papineau, D. (2001). The rise of physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and its discontents* (pp. 3–36). Cambridge University Press.

Papineau, D. (2002). *Thinking about consciousness*. Clarendon Press.

Papineau, D. (2006). Comments on Galen Strawson. In A. Freeman (Ed.), *Consciousness and its place in nature* (pp. 3–31). Imprint Academic.

Papineau, D. (2008). Explanatory gap and dualist intuitions. In L. Weiskrantz & M. Davies (Eds.), *Frontiers of consciousness* (pp. 55–68). Oxford University Press.

Papineau, D. (2020). The problem of consciousness. In U. Kriegel (Ed.), *The Oxford handbook of the philosophy of consciousness* (pp. 14–38). Oxford University Press.

Pasnau, R. (2011). *Metaphysical themes: 1274–1671*. Clarendon Press.

Pautz, A. (2013). Does phenomenology ground mental content? In U. Kriegel (Ed.), *Phenomenal intentionality: New essays* (pp. 194–234). Oxford University Press.

Pautz, A. (2017). The significance argument for the irreducibility of consciousness. *Philosophical Perspectives*, *31*(1), 349–407.

Penfield, W. (1975). *Mystery of the mind: A critical study of consciousness and the human brain*. Princeton University Press.

Pereira, A., Jr., & Ricke, H. (2009). What is consciousness? Towards a preliminary definition. *Journal of Consciousness Studies*, *16*, 28–45.

Peressini, A. F. (2018). There is nothing it is like to see red: Holism and subjective experience. *Synthese*, *195*, 4637–4666.

Perler, D. (2015). *The faculties: A history*. Oxford University Press.

Perry, J. (2001). *Knowledge, possibility, and consciousness*. MIT Press.

Peter, V. I. (1990). *Material beings*. Cornell University Press.

Peter, V. I. (1995). Dualism and materialism: Athens and Jerusalem? *Faith and Philosophy*, *12*(4), 475–488.

Peter, V. I. (1998). Modal epistemology. *Philosophical Studies*, *92*, 67–84.

Peter, V. I. (2004). The self: The incredulous stare articulated. *Ratio*, *17*(4), 478–491.

Peter, V. I. (2006). Can mereological sums change their parts? *Journal of Philosophy*, *103*(12), 614–630.

Peter, V. I. (2007). Plantinga's replacement argument. In D.-P. Baker (Ed.), *Alvin Plantinga*. Oxford University Press.

Peter, V. I. (2014). *Existence: Essays in ontology*. Cambridge University Press.

Peters, E. (2006). Numeracy and decision making. *Psychological Science*, *17*, 407–413.

Peters, E. P. S., & Gregory, R. (2003). The role of affect in the WTA/WTP disparity. *Journal of Behavioral Decision Making*, *16*, 309–330.

Pinker, S. (1997). *How the mind works*. W.W. Norton.

Pinto, Y. (2017a). Split brain: Divided perception but undivided consciousness. *Brain: A Journal of Neurology*, *140*(5), 1231–1237.

Pinto, Y. (2017b). The split-brain phenomenon revisited: A single conscious agent with split perception. *Trends in Cognitive Sciences*, *21*(11), 835–851.

Pitt, D. (2004). The phenomenology of cognition or what is it like to think that P? *Philosophy and Phenomenological Research*, *LXIX*(July), 1–36.

Pitt, D. (2019). Acquaintance and phenomenal concepts. In S. Coleman (Ed.), *The knowledge argument* (pp. 87–101). Cambridge University Press.

Pitts, J. B. (2022). General relativity, mental causation, and energy conservation. *Erkenntnis*, *87*. https://doi.org/10.1007/s10670-020-00284-7

Place, U. T. (2002). A Pilgrim's progress? From mystical experience to biological consciousness. *Journal of Consciousness Studies*, *9*, 34–52.

Plantinga, A. (1993). *Warrant and proper function*. Oxford University Press.

Plantinga, A. (2007). Materialism and Christian belief. In V. I. Peter & D. Zimmerman (Eds.), *Persons: Human and divine*. Oxford University Press.

Plato. (1997). *Plato: Complete works* (J. M. Cooper, Ed., with Introduction & Notes). Hackett.

Pollock, J. L. (1974). *Knowledge and justification*. Princeton University Press.

Popper, K. R., & Eccles, J. C. (1985). *The self and its brain*. Springer-Verlag.

Potts, M. (2001). Sensory experiences in near death experiences and the Thomistic view of the soul. *International Journal for Philosophy of Religion*, *49*, 85–100.

Preston, A. (2015). Intuition in analytic philosophy. *Teorema*, *34*(3), 37–56.

Preston, J. L., Ritter, R., & Hepler, J. (2013). Neuroscience and the soul: Competing explanations for the human experience. *Cognition*, *127*, 31–37.

Pruss, A. R. (2011). *Actuality, possibility and worlds*. Bloomsbury.

Pust, J. (2000). *Intuitions as evidence*. Routledge.

Putnam, H. (1963). Brains and behavior. Mind, Language, and Reality. *Philosophical Papers*, *2*, 325–341.

Putnam, H. (1975). Philosophy and our mental life. In H. Putman (Ed.), *Mind, language, and reality*. Cambridge University Press.

Putnam, H. (1981). *Reason, truth, and history*. Cambridge University Press.

Pyysiäinen, I. (2009). *Supernatural agents: Why we believe in souls, gods, and Buddhas*. Oxford University Press.

Rachels, J. (1990). *Created from animals*. University Press.

Raleigh, T. (2019). Introduction: The recent renaissance of acquaintance. In J. Knowles & T. Raleigh (Eds.), *Acquaintance: New essays* (pp. 1–33). Oxford University Press.

Raleigh, T., & Knowles, J. (Eds.). (2019). *Acquaintance: New essays*. Oxford University Press.

Rancurello, A. C. (1968). *A study of Franz Brentano: His psychological standpoint and his significance in the history of psychology*. Academic Press.

Rasmussen, J. (2018). Against nonreductive physicalism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 336–337). Wiley-Blackwell.

Ravenscroft, I. (2005). *Philosophy of mind: A beginner's guide*. Oxford University Press.

Ravenscroft, I. (2018). Why reject substance dualism? In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism*. Wiley-Blackwell.

Rea, M. (2011). Hylomorphism reconditioned. *Philosophical Perspectives*, *25*, 341–358.

Rea, M. C. (2002). *World without design: The ontological consequences of naturalism*. Oxford University Press.

Reed, E. S. (1997). *From soul to mind*. Yale University Press.

Reid, T. (1983). Of identity. In R. E. Beanblossom & K. Lehrer (Eds.), *Thomas Reid: Inquiry and essays* (pp. 214–215). Hackett.

Richards, J. W. (Ed.). (2002). *Are we spiritual machines? Ray Kurzweil vs. the critiques of strong AI*. Discovery Institute.

Rickabaugh, B. (2013). Eternal life as knowledge of God: An epistemology of knowledge by acquaintance and spiritual formation. *Journal of Spiritual Formation and Soul Care*, *6*(2), 204–228.

Rickabaugh, B. (2018). Against emergent dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism* (pp. 73–86). Wiley-Blackwell.

Rickabaugh, B. (2019). Alister McGrath's anti-mind-body dualism: Neuroscientific and philosophical quandaries for Christian physicalism. *Trinity Journal*, *40*, 215–240.

Rickabaugh, B. (2020). *The conscious mind unified*. Baylor University.

Rickabaugh, B. (work in preparation). *Consciousness and fundamental fine-tuning: Franz Brentano's teleological insights contra agentive cosmopsychism*.

Rickabaugh, B. (work in preparation). *Epistemology, mind, and hard choices: Externalism or dualism, but not both*.

Rickabaugh, B., & Stephen Evans, C. (2018). Neuroscience, spiritual formation, and bodily souls: A critique of christian physicalism. In R. K. Loftin & J. Farris (Eds.), *Christian physicalism? Philosophical theological criticisms* (pp. 231–256). Lexington Books.

Rickabaugh, B. (work in preparation). *Uncovering Thomas Reid's hylomorphic substance dualism*.

Rickabaugh, B. (work in preparation). *Enduring problems for animalism*.

Roazzi, M., Nyhof, M., & Johnson, C. (2013). Mind, soul and spirit: Conceptions of immaterial identity in different cultures. *International Journal for the Psychology of Religion*, *23*, 75–86.

Robert, V. G. (2007). Who's in charge here? And who's doing all the work? In N. Murphy & W. R. Stoeger (Eds.), *Evolution and emergence: Systems, organisms, persons*. Oxford University Press.

Robinson, H. (1987). A dualist perspective on psychological development. In J. A. Russell (Ed.), *Philosophical perspectives on developmental psychology* (pp. 119–139). Basil Blackwell.

Robinson, H. (1989). A dualist account of embodiment. In J. R. Smythies & J. Beloff (Eds.), *The case for dualism* (pp. 43–57). University of Virginia Press.

Robinson, H. (1993). The anti-materialist strategy and the 'knowledge argument'. In H. Robinson (Ed.), *Objections to physicalism* (pp. 159–184). Clarendon.

Robinson, H. (2010). Substance dualism and its rationale. In R. Swinburne (Ed.), *Free will and modern science* (pp. 158–177). Oxford University Press.

Robinson, H. (2016). *From the knowledge argument to mental substance: Resurrecting the mind*. Cambridge University Press.

Robinson, H. (2021). The revival of substance dualism. *Roczniki Filozoficzne*, *LXIX*(1), 33–43.

Robinson, W. (2004). *Understanding phenomenal consciousness*. Cambridge University Press.

Rochat, P., & Striano, T. (1999). Social-cognitive development in the first year. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first year of life* (pp. 3–34). Lawrence Erlbaum.

Rockwell, T. (2005). *Neither ghost nor brain*. MIT Press.

Rodrigues, J. G. (2014). There are no good objections to substance dualism. *Philosophy*, *89*(2), 199–222.

Roelofs, L. (2016). The unity of consciousness, within and between subjects. *Philosophical Studies*, *173*, 3199–3221.

Rojka, Ľ. (2016). The modal argument for soul/body dualism. *Studia Neoaristotelica*, *13*(1), 45–70.

Rosen, G., & Dorr, C. (2002). Composition as a fiction. In R. M. Gale (Ed.), *The Blackwell guide to metaphysics*. Blackwell.

Rosenberg, G. H. (2004). *A place for consciousness: Probing the deep structure of the natural world*. Oxford University Press.

Rosiak, M. (2017). Edmund Husserl. In J. Seibt, G. Imaguire, & S. Gerogiorgakis (Eds.), *Handbook of mereology (München)* (pp. 262–269). Philosophia Verlag.

Rozemond, M. (1998). *Descartes's dualism*. Harvard University Press.

Rozemond, M. (2003). Descartes, mind-body union, and holenmerism. *Philosophical Topics*, *31*(1/2), 343–367.

Russell, B. (1911). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, New Series, *11*(1910–1911), 108–128.

Russell, B. (1914). On the nature of acquaintance. *Monist*, *24*, 435–453.

Russell, B. (1948). *Human knowledge: Its scope and limits*. Simon and Schuster.

Russell, B. (2001). *Problems of philosophy* (2nd ed.). Oxford University Press.

Russell, M. (2006). *Husserl: A guide for the perplexed*. Continuum.

Saad, B. (2018). A causal argument for dualism. *Philosophical Studies*, *175*(10), 2475–2506.

Sabatés, M. H. (2005). Reductionism in the philosophy of mind. *Encyclopedia of Philosophy*. http://www.encyclopedia.com/humanities/encyclopedias-almanacs-transcripts-and-maps/reductionism-philosophy-mind

Salice, A. (2017). Franz Brentano. In J. Seibt, G. Imaguire, & S. Gerogiorgakis (Eds.), *Handbook of mereology (München)* (pp. 119–125). Philosophia Verlag.

Samuels, R. (2009). The magical number two, plus or minus: Dual process theory as a theory of cognitive kinds. In K. Frankish & J. S. B. T. Evans (Eds.), *Two minds: Dual processes and beyond* (pp. 129–146). Oxford University Press.

Sandberg, A., & Bostrom, N. (2008). Whole brain emulation: A roadmap" *Technical Report #2008-3*. Future of Humanity Institute, Oxford University.

Sarkissian, H., Chatterjee, A., de Brigard, F., Knobe, J., Nichols, S., & Sirker, S. (2010). Is belief in fee will a cultural universal? *Mind and Language*, *25*, 346–358.

Satel, S., & Lilienfel, S. (2013). *Brainwashed: The seductive appeal of mindless neuroscience*. Basic Books.

Saxe, R., Tzelnic, R., & Carey, S. (2006). Five-month-old infants know humans are solid, like inanimate objects. *Cognition*, *101*(1), 1–8.

Schärtl, T. (2012). Beyond dualism? The track-switch model of resurrection. In B. P. Göcke (Ed.), *After physicalism* (pp. 335–368). University of Notre Dame Press.

Schechter, E. (2018). The unity of consciousness. In R. J. Gennaro (Ed.), *The Routledge handbook of consciousness* (pp. 366–378). Routledge.

Schneider, S. (2012). Why property dualists must reject substance physicalism. *Philosophical Studies*, *157*(1), 61–76.

Schneider, S. (2013). Non-reductive physicalism and the mind problem. *Noûs*, *47*(1), 135–153.

Schneider, S. (2019). *Artificial you: AI and the future of your mind*. Princeton University Press.

Schneider, S. (2020). How to catch an AI Zombie: Testing for consciousness in machines. In S. Matthew Liao (Ed.), *Ethics of artificial intelligence* (pp. 439–458). Oxford University Press.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and 'theory of mind'. *Mind and Language*, *14*, 131–153.

Schouten, M. K. (2001). Theism, dualism, and the scientific image of humanity. *Zygon*, *36*(4), 679–708.

Schwartz, J. M., Stapp, H. P., & Beauregard, M. (2005). Quantum physics in neuroscience and psychology: A neurophysical model of mind—brain interaction. *Philosophical Transactions of the Royal Society Biological Sciences*, *360*, 1309–1327.

Schwartz, J. M., & Begley, S. (2002). *The mind and the brain: Neuroplasticity and the power of mental force*. HarperCollins.

Schwitzgebel, E. (2007). Do you have constant tactile experience of your feet in your shoes? Or is experience limited to what's in attention? *Journal of Consciousness Studies*, *14*, 5–35.

Scruton, R. (2017). If we are not just animals, what are we? *New York Times*, March 6. https://mobile.nytimes.com/2017/03/06/opinion/if-we-are-not-just-animals-what-are-we.html?referer=http%3A%2F%2Fwww.thepublicdiscourse.com%2F2017%2F05%2F19297%2F

Seager, W. (2010). Panpsychism, aggregation and combinatorial infusion. *Mind and Matter*, *8*(2), 167–184.

Searle, J. (2002). Why I am not a property dualist. *Journal of Consciousness Studies*, *9*(12), 57–64.

Searle, J. (2007). *Freedom and neurobiology*. Colombia University Press.

Searle, J. R. (1980a). Intrinsic intentionality. *Behavioral and Brain Sciences*, *3*, 450–456.

Searle, J. R. (1980b). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417–457.

Searle, J. R. (1984). *Minds. brains, and science*. Harvard University Press.

Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.

Searle, J. R. (2004). *Mind: A brief introduction*. Oxford University Press.

Seligman, M., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science*, *8*, 119–141.

Sellars, W. (1956). Empiricism and the philosophy of mind. In H. Feigl & M. Scriven (Eds.), *Minnesota studies in the philosophy of science* (Vol. I, pp. 127–196). University of Minnesota Press.

Sellars, W. (1963). *Science, perception and reality*. Routledge and Kegan Paul.

Shafer-Landau, R. (2003). *Moral realism: A defense*. Clarendon Press.

Shrader, W. (2006). The unity of consciousness: Trouble for the materialist or the emergent dualist? *Faith and Philosophy*, *23*(1), 33–44.

Shulman, R. (2013). *Brain imaging: What it can (and cannot) tell us about consciousness*. Oxford University Press.

Sider, T. (2001). Maximality and intrinsic properties. *Philosophy and Phenomenological Research*, *63*, 357–364.

Sider, T. (2003). Maximality and microphysical supervenience. *Philosophy and Phenomenological Research*, *66*, 139–149.

Siegel, D. J. (2010). *Mindsight: The new science of personal transformation*. Random House.

Siewert, C. (1998). *The significance of consciousness*. Princeton University Press.

Simpson, W. M. R. (2021). Cosmic hylomorphism. *European Journal for Philosophy of Science*, *11*(1), 1–25.

Simpson, W. M. R. (2022). From quantum physics to classical mechanics. In W. M. R. Simpson, R. C. Koons, & J. Orr (Eds.), *Neo-Aristotelian metaphysics and the theology of nature* (pp. 21–65). Routledge.

Simpson, W. M. R., Koons, R. C., & Teh, N. J. (Eds.). (2018). *Neo-Aristotelian perspectives on contemporary science*. Routledge.

Skirry, J. (2001). A hylomorphic interpretation of Descartes's theory of mind-body union. *Proceedings of the American Catholic Philosophical Association*, *75*, 267–283.

Skirry, J. (2005). *Descartes and the metaphysics of human nature*. Thoemmes-Continuum Press.

Skrbina, D. (2005). *Panpsychism in the west*. MIT Press.

Sławkowski-Rode, M. (2021). The distinction in question: The analytic/continental divide in philosophy. In B. P. Göcke & R. S. Weir (Eds.), *From existentialism to metaphysics* (pp. 45–62). Peter Lang.

Slingerland, E., & Chudek, M. (2011). The prevalence of mind-body dualism in early China. *Cognitive Science*, *35*(5), 997–1007.

Smith, B., & Mulligan, K. (1982). Pieces of a theory. In B. Smith (Ed.), *Parts and moments: Studies in logic and formal ontology* (pp. 15–109). Philosophia Verlag.

Smith, D. W., & McIntyre, R. (1982). *Husserl and intentionality: A study of mind, meaning, and language*. D. Reidel.

Smithies, D. (2019). *The epistemic role of consciousness*. Oxford University Press.

Smolin, L. (2006). *The trouble with physics: The rise of string theory, the fall of a science and what comes next*. A Mariner Book.

Sokolowski, R. (1968). The logic of parts and wholes in Husserl's investigations. *Philosophy and Phenomenological Research*, *28*(4), 537–553.

Solère, J.-L. (2017). Bayle and panpsychism. *Archiv für Geschichte der Philosophie*, *99*(1), 64–101.

Sosa, E. (1984). Mind-body interaction and supervenient causation. *Midwest Studies in Philosophy*, *9*, 271–281.

Sosa, E. (2007). Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien*, *74*, 51–67.

Spelke, E. S. (1996). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44–78). Clarendon Press.

Springer, S. P., & Deutsch, G. (1981/1998). *Left brain right brain: Perspectives from cognitive neuroscience*. W. H. Freeman/Worth.

Stafford, B. L. (2004). Mind, paranormal experience, and the inadequacy of materialism. *International Philosophical Quarterly*, *44*(3), 373–392.

Stainton, R. J. (Ed.). (2006). *Contemporary cognitive science*. Blackwell.

Steinhart, E. (2019). Naturalism. In G. Oppy (Ed.), *A companion to atheism and philosophy* (pp. 152–166). Wiley-Blackwell.

Stephan, E. C. (2005). Separable souls: dualism, selfhood, and the possibility of life after death. *Christian Scholar's Review*, *34*, 327–340.

Stephen, E. C., & Rickabaugh, B. (2015). What does it mean to be a bodily soul? *Philosophia Christi*, *17*(2), 315–330.

Steward, H. (2012). *A metaphysics for freedom*. University Press.

Stoljar, D. (2010). *Physicalism*. Routledge.

Strawson, G. (1997). The self. *Journal of Consciousness Studies*, *4*(5–6), 405–428.

Strawson, G. (2006). Realistic monism: Why physicalism entails panpsychism. In A. Freeman (Ed.), *Consciousness and its place in nature* (pp. 3–31). Imprint Academic.

Strawson, G. (2009). *Selves: An essay in revisionary metaphysics*. Oxford University Press.

Strawson, G. (2010). Fundamental singleness: How to turn the 2nd paralogism into a valid argument. In P. Basile, J. Kiverstein, & P. Phemister (Eds.), *The metaphysics of consciousness 67, Royal Institute of Philosophy Supplement* (pp. 61–92). Cambridge University Press.

Strawson, G. (2017). *The subject of experience*. Oxford University Press.

Stump, E. (2010). *Wandering in darkness: Narrative and the problem of suffering*. Clarendon Press.

Stump, E. (2013). Emergence, causal powers and Aristotelianism in metaphysics. In R. Groff & J. Greco (Eds.), *Powers and capacities: The new Aristotelianism* (pp. 48–68). Routledge.

Suárez, F. (2000). *On the formal cause of substance: Metaphysical disputations XV* (J. Kronen & J. Reedy, Trans.). Marquette University Press.

Swinburne, R. (1997a). *The evolution of the soul* (Revised ed.). Clarendon Press.

Swinburne, R. (1997b). *Simplicity as evidence of truth*. Marquette University Press.

Swinburne, R. (2001). *Epistemic justification*. Oxford University Press.

Swinburne, R. (2003). What is so good about having a body? In T. W. Bartel (Ed.), *Comparative theology: Essays for Keith Ward* (pp. 134–142). SPCK Publishing.

Swinburne, R. (2013). *Mind, brain, and free will*. University Press.

Swinburne, R. (2018). *The argument from colors and flavors: The argument from consciousness* (J. L. Walls & T. Dougherty, Eds.) Two dozen (or so) arguments for god (pp. 293-303). Oxford University Press.

Tabaczek, M. (2019). *Emergence*. Notre Dame Press.

Taliaferro, C. (1986). A modal argument for dualism. *Southern Journal of Philosophy*, *24*(1), 95–108.

Taliaferro, C. (1994). *Consciousness and the mind of god*. Oxford University Press.

Taliaferro, C. (1995). Animals, brains, and spirits. *Faith and Philosophy*, *12*(4), 567–581.

Taliaferro, C. (2001a). *Emergentism and consciousness* (K. Corcoran, Ed.). Cornell University Press.

Taliaferro, C. (2001b). The virtues of embodiment. *Philosophy*, *76*, 111–125.

Taliaferro, C. (2014). The promise and sensibility of integrative dualism. In A. Lavazza & H. Robinson (Eds.), *Contemporary dualism: A defense* (pp. 191–203). Routledge.

Taliaferro, C. (2017). Dualism and panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary perspectives* (pp. 369–387). Oxford University Press.

Taliaferro, C. (2018). Substance dualism: A defense. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell companion to substance dualism*. Blackwell.

Taliaferro, C. (2021). In defense of integrative dualism; Placing values at the heart of philosophy of mind. *Theosophia Islamica*, *1*(1), 33–55.

Tallis, R. (2011). *Aping mankind: Neuromania, Darwinitis and the misrepresentation of humanity*. Routledge.

Taylor, C. (1989). *Sources of the self: The making of the modern identity*. Cambridge University Press.

Taylor, H. (2020). The relation between subjects and their conscious experiences. *Philosophical Studies*, *177*(11), 3493–3507.

Taylor, R. (1997). Chisholm's idea of a person. In L. E. Hahn (Ed.), *The philosophy of Roderick M. Chisholm*. Open Court.

Textor, M. (2021). *The disappearance of the soul and the turn against metaphysics: Austrian philosophy 1874–1918*. Oxford University Press.

Thornton, A. K. (2016). Varieties of animalism. *Philosophy Compass*, *11*(9), 515–526.

Thurow, J. C. (2018). Animals with souls. *Sophia*, *57*(1), 85–101.

Tidman, P. (1994). Conceivability as a test for possibility. *American Philosophical Quarterly*, *31*(4), 297–308.

Tiller, P. (2016). Supervenience. In K. J. Clark (Ed.), *The Blackwell companion to naturalism*. John Wiley.

Timpe, K., & Jacobs, J. (2016). Free will and naturalism: How to be a libertarian, and a naturalist, too. In K. J. Clark (Ed.), *The Blackwell companion to naturalism*. John Wiley.

Toepfer, G. (2012). Teleology and its constitutive role for biology as the science of organized systems in nature. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 113–119.

Toner, P. (2011a). Hylemorphic animalism. *Philosophical Studies*, *155*(1), 65–81.

Toner, P. (2011b). Independence accounts of substance and substantial parts. *Philosophical Studies*, *155*(1), 37–43.

Tooley, M. (2009). The appeal to an immaterial mind. In M. Tooley, C. Wolf-Devine, P. E. Devine, & A. Jaggar (Eds.), *Abortion: Three perspectives* (Vol. 19). Oxford University Press.

Unger, P. (2004). The mental problems of the many. In D. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. I, pp. 196–222). Clarendon Press.

Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. MIT Press.

Uttal, W. R. (2004). *Dualism: The original sin of cognitivism*. Lawrence Erlbaum Associates.

Vandergriff, K. (2018). Naturalism, theism, and multiply realizable mental states. *Religious Studies*, *54*, 91–105.

Vargas, M. (2014). Reconsidering scientific threats to free will. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 4: *Free Will and Moral Responsibility*, pp. 417–425). MIT Press.

Velarde-Mayol, V. (2000). *On Husserl*. Wadsworth.

Velmans, M. (2007). Where experiences are: Dualist, physicalist, enactive and reflexive accounts of phenomenal consciousness. *Phenomenology and Cognitive Science*, *6*, 547–563.

Velmans, M. (2009). *Understanding consciousness* (2nd ed.). Routledge.

Ventriglio, A., & Bhugra, D. (2015). Descartes' dogma and damage to western psychiatry. *Epidemiology and Psychiatric Sciences*, *5*, 368–370.

Vimal, R. L. P. (2009). Meanings attributed to the term 'consciousness': An overview. *Journal of Consciousness Studies*, *16*(5), 9–27.

Voss, S. (1992). Understanding eternal life. *Faith and Philosophy*, *9*(1), 3–22.

Wald, R. M. (1984). *General relativity*. University of Chicago Press.

Webster, G., & Goodwin, B. (1996). *Form and transformation: Generative and relational principles in biology*. Cambridge University Press.

Weir, R. S. (2021a). Bring back substances! *The Review of Metaphysics*, *75*(2), 265–308.

Weir, R. S. (2021b). Can a post-Galilean science of consciousness avoid substance dualism? *Journal of Consciousness Studies*, *28*(9–10), 212–228.

Weir, R. S. (2023). *Metaphysics and the mind-body problem: Why the twenty-first century still needs nonphysical substances*. Routledge.

Wellman, H. (1990). *The child's theory of mind*. MIT Press.

Wellman, H., & Estes, D. (1986). Early understanding of mental entities: A re-examination of childhood realism. *Child Development*, *57*, 910–923.

Wellman, H., & Hickling, A. K. (1994). The minds 'I': Children's conception of the mind as an active agent. *Child Development*, *65*, 1564–1580.

Welsh, T. (2015). It feels instantaneous, but how long does it really take to think a thought? *The Conversation*, June 26. http://theconversation.com

Westerhoff, J. (2005). *Ontological categories: Their nature and significance*. Clarendon Press.

Whatley, S. D. (2014). Borrowed philosophy: Bedside physicalism and the need for a sui generis metaphysic of medicine. *Journal of Evaluation in Clinical Practice*, *6*, 961–974.

White, R., & Grünbaum, A. (2005). Why favour simplicity? *Analysis*, *65*(3), 205–210.

Wiggins, D. (2001). *Sameness and substance*. University Press.

Willard, D. (1982). Wholes, parts and the objectivity of knowledge. In B. Smith (Ed.), *Parts and moments: Studies in logic and formal ontology*. Philosophia Verlag.

Willard, D. (1984). *Logic and the objectivity of knowledge: A study in Husserl's early philosophy*. Ohio University Press.

Willard, D. (1989/2017). Who needs Brentano? The wasteland of philosophy without its past. In R. Poli (Ed.), *The Brentano puzzle* (pp. 15–43). Routledge.

Willard, D. (1994). Mereological essentialism restricted. *Axiomathes*, *5*(1), 123–144.

Willard, D. (1995). Knowledge. In B. Smith & D. W. Smith (Eds.), *The Cambridge companion to Husserl* (pp. 138–167). Cambridge University Press.

Willard, D. (1999). How concepts relate the mind to its objects: The God's eye view vindicated. *Philosophia Christi*, *1*(2), 5–20.

Willard, D. (2000). Knowledge and naturalism. In W. L. Craig & J. P. Moreland (Eds.), *Naturalism: A critical analysis* (pp. 24–48). Routledge.

Willard, D., & Rickabaugh, B. (2018). Intentionality contra physicalism: On the mind's independence from the body. *Philosophia Christi*, *20*(2), 497–515.

Wilson, J. (2010). Review of *The mind in nature*. *Mind*, *119*, 503–511.

Wiltsche, H. (2015). Intuitions, seemings, and phenomenology. *Teorema*, *34*(3), 57–78.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Wimsatt, W. (1995). The ontology of complex systems. *Canadian Journal of Philosophy*, *20*, 564–590.

Wisniewski, D., Deutschländer, R., & Haynes, D. (2019). Free will beliefs are better predicted by dualism than determinism beliefs across different cultures. *PLOS ONE*, *14*(9), 1–19.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, & R. Rhees, Eds.). Translated by G.E.M. Anscombe. Oxford University Press.

Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, *53*(1), 1–42.

Yandell, K. (1995). A defense of dualism. *Faith and Philosophy*, *12*(4), 548–566.

Yang, E. (2018). Persons, simplicity, and substance. *Philosophical Papers*, *47*(2), 299–311.

Yolton, J. (1983). *Thinking matter: Materialism in eighteenth-century Britain*. University of Minnesota Press.

Zagzebski, L. T. (2021). *The two greatest ideas: How our grasp of the universe and our minds changed everything*. Princeton University Press.

Zerin, E. (1998). Karl Popper on god: The lost interview. *Skeptic*, *6*(2), 47–48.

Zimmerman, D. (1998). Immanent causation. In J. E. Tomberlin (Ed.), *Philosophical perspectives. Vol. 11: Mind, causation, and world*. Ridgeview.

Zimmerman, D. (2006). Dualism in the philosophy of mind. In D. M. Borchert (Ed.), *Encyclopedia of philosophy* (2nd ed., Vol. 3, pp. 113–122). Thomson Gale.

Zimmerman, D. (2010). From property dualism to substance dualism. *Aristotelian Society Supplementary*, *84*(1), 119–150.

Zimmerman, D. (2011). From experience to experiencer. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis; Investigations into the existence of the soul* (pp. 168–196). Continuum.

Index

Page numbers in *italic* indicate figures; page numbers followed by 'n' indicate footnotes. The 'SD' (substance dualism) abbreviation is used throughout the index.

A
abstractness, concept of, 210
absurd introspection skepticism, 63–6
ACT *see* AI Consciousness Test
active power
  characteristics of, 259, 260
  and conservation of energy principle, 279
  denial of, 248
  free agents, 243
  top-down causation, 247
active thoughts, 100
AFI (argument from introspection) *see* introspection, argument from
afterlife, 88, 107, 341
AGENCY account, cognitive processes, 111, 112–13, 114
agent causation, 234
aggregates, 128
  Husserlian ontology, 46–50, *49*, 56
  macro-objects as, 266–7
AI (artificial intelligence), 340–2
AI Consciousness Test (ACT), 90
akrasia ("weakness of will"), 239
analytic phenomenology, 322
ancient world, and substance dualism, 91
animal composition, 130
animalism, 9, 129–31, 204
anti-distribution argument, 123–4
applied SD research projects, 332–42
  content externalism, 336
  epistemology, 334–5
  healthcare and mental health, 337–9
  philosophy of religion, 332–4
  social psychology, 337
  technology, 339–42
a priori knowledge, 213, 221
argument from introspection (AFI) *see* introspection, argument from
Aristotelian atom, 152
Aristotelianism, revival of, 8–9
Aristotle, 92, 105, 246
  see also Medieval Aristotelians; Metaphysical Aristotelianism
Armstrong, D. M., 46, 148
artificial intelligence (AI), 340–2
artificial intelligence test seemings, 90
atheism, 7, 8, 39
atomic simples, human persons as, 169–72
atomless gunk, 152
atoms, 40
Austin, Christopher J., 172

B
Bailey, Andrew, 22, 76, 150
  contingent physicalism, 201–12
  impossibility argument, objection to, 305–7
  parsimony argument, 300–1
Bandettini, Peter, 52
Bare Particular View, 146
Barnett, David, philosophy of mind, 85–6
basicality, of phenomena, 20
Basile, P., 139
Bayne, Tim, 285
  atomistic theory of consciousness, 121
  subject phenomenal unity, 122
  unity of consciousness, explanatory failure objection, 142–3
Bealer, George
  a priori knowledge, 213
  modal knowledge, 218–23

The Substance of Consciousness: A Comprehensive Defense of Contemporary Substance Dualism, First Edition. Brandon Rickabaugh and J. P. Moreland.
© 2024 John Wiley & Sons, Inc. Published 2024 by John Wiley & Sons, Inc.

Bedau, Mark, 31, 41
Begley, Sharon, 331, 349–50
beliefs
  belief-experiences, 101
  commonsense beliefs, 74
  de dicto and *de re* beliefs, 62–3, 67–72, 231
  and intuition, 109–10
  modal beliefs, 213, 215
  wishful beliefs, 350
Bennett, Karen, 276
Bergmann, Gustav, 46
Bering, Jesse, 87–8
Big Bang theory, 27, 281
Bishop, John, 235
  free will, compatibilist theory, 239–40
  natural causal relations, 242–3
Bjorklund, David, 87–8
Block's Miniature Men in the Head, 85
Bloom, Paul, 86
bodily soul view, 325–6
the body
  accidental immateriality, 198
  Metaphysical Aristotelianism, 183
  and soul, 306, 307, 345
  tethered body/tethered soul, 306, 307
  transhumanist view of, 340
  virtual body, 342
  see also embodiment
brain
  bisection surgery, 141
  dispositions, 342
  entropic brain hypothesis, 350–1
  imaging techniques, 347–50
  physically uncaused spikes in energy level of (undetected), 346–57
  quantum stimuli, responsiveness to, 357–9
  soul interactions, 344–60
  and thinking, 295
brain states
  and consciousness, 114, 248
  explanatory gap, 255
  and mental states, 67–8, 247–8, 359
  split-brain phenomena, 133, 141
Brandt, Richard, 59–60
Broad, C. D., 280
broken window analogy, causation, 265–6
Broks, Paul, 86
brute facts, emergence as, 37–8
Buddhism, 148

C

CA *see* causal-acquaintance principle
Cain, M. J., 347
Campbell, John, 196, 214
Campbell, Keith, 280
capacities/faculties, metaphysical categorization of, 287–8
car engine analogy, functional roles, 207
Carhart-Harris, Robin L., 350–1
Cartesian Commitments, 312, 313
Cartesian dualism, 314, 315, 353
Cartwright, S., 326
categorical ability, 245
categorical contingency, 208
categorical intuition, 225
categorical theory, 342
causal-acquaintance principle (CA), 100
causal chain analysis, 214–15, 217, 218
causal closure
  libertarian free will, 263–70
  of physical domain, 248, 256, 353, 354, 355, 356
causal interaction, problems of, 276–83
  causal pairing problem, 280–3
  conservation of energy principle, violation of, 277–80
  mental/physical entity causal interaction, 276–7
causality/causation
  causal-powers view of, 263, 264, 266, 267, 269
  complex whole causation, 265–6
  consciousness, 257
  downward causation, 177–8
  event-causation, 28, 242
  event–event causation, 241, 293
  free will, 259–62
  immanent causation, 243, 244
  indeterminate causation, 260
  mental causation, *247*
  mental properties, 268–9
  mereological hierarchy, 153
  natural causal relations, 242–3
  naturalistic causation, 244
  passive liabilities, 243
  state–state causation, 293
  substance causation, 242, 265
  top-down causation, 153, 154, 246–9, 256, 257, 263–70
causal likeness principle, 277
causal modal knowledge, 215
causal overdetermination (COD), arguments against, 265–7
causal pairing, problem of, 280–3
Chalmers, David, 6, 9–11, 13, 102–3
  consciousness, 82–3, 354–6
  Palette Problem, 38–9
  panpsychism, 134
  subject phenomenal unity, 122
chemical synapses, 81
children, understanding of mind study, 87

- Chisholm, Roderick, 98, 100, 106, 295
  - extended physical atom view, 169–72
  - unified consciousness, 125
- Churchland, Patricia, first philosophy, rejection of, 25, 291, 292
- Churchland, Paul, 26, 33, 157
  - argument from introspection, objections to, 60–72
  - dualism/physicalism debate, adjudication criteria, 286–7
  - scanning-monitoring approach, 115–16
- circularity problem, personal identity, 165–6
- closure principle *see* causal closure
- C-maximality, 78, 207
- COD *see* causal overdetermination
- cognitive pathways, 111–12
- cognitive processes, AGENCY features, 111, 112–13, 114
- cognitive science
  - commonsense substance dualism, 86–8
  - consciousness and, 8
  - explanatory gap seemings, 110–11
- Collins, R., 278
- color sensation, 99
- combination problem, 38–9, 134–5
- commonsense beliefs, 74, 86–8, 240
- commonsense materialism, 106
- compatibilist theory, free will, 239–40
- compatibility/incompatibility, objects, 228–9
- complex objects, proper constituents, 198
- complex systems
  - as enduring continuants, 178–9
  - as synchronically unified wholes, 175–7
- complex whole causation (CWC), 265–6
- compound dualism, 322
- conceivability, modal argument for SD, 212–13
- concepts and laws, mereological hierarchy, 153–5
- conceptual SD research projects, 312–16
- confirmation theory, likelihood principle, 296
- conjoined twins, mental-sharing, 140
- conscious cup, 80–1
- consciousness
  - anthropic fine-tuning debates, 334
  - atomistic theory of, 121
  - Barnett, 75–82
  - and brain states, 114, 248
  - Carhart-Harris, 350–1
  - causality, 257
  - C-maximality, 78, 207
  - cognitive pathways, 111–12
  - and cognitive science, 8
  - combinatorial modes of explanation, 25–6
  - consciousness-collapse models, 354–6
  - continuous flow of, 65
  - defining by ostension, 50–3
  - Dennett, 93
  - easy meta-problem of, 83
  - emergence of, 39–40, 115, 252, 255, 256
  - entropic brain hypothesis, 350–1
  - first-person approach, 50–3, 293
  - fragmentation of, 133
  - and free will, 252–4
  - fundamentality of, 10
  - hard meta-problem of, 82–92
  - Hopp, 286
  - integrity, 79–81
  - and material/immaterial substance, 302
  - as maximal property, 77–8
  - measurement of, 289
  - mental states and, 55
  - metaphysics of, 10, 50–6
  - mine-ness, 88
  - neural correlates of, 289
  - neurobiological level of, 253, 254, 257, 284
  - origin of, 36, 307
  - phenomenal unity, 89
  - and philosophy of mind, 11
  - primary/secondary consciousness distinction, 350
  - primitive consciousness, 96
  - psychedelic drugs, 350, 351
  - and the self, 8, 53–4, 89
  - simplicity argument, Barnett's, 75–82
  - soul, explanatory impotence of, 301–7
  - subject necessity, 54–5, 88
  - of subjects, 55
  - subject simplicity and complexity, 56
  - and substance, 78–9
  - substantial subjects and the self, realism about, 53–4
  - tethered body/tethered soul, 306, 307
  - theories of, 33–4
  - and wave function collapse, 354–6
  - *see also* phenomenal unity of consciousness
- consciousness-first approach, 320–2, 333
- conservation of energy principle, and causal interaction, 277–80
- conservation laws, 277–80, 352, 353, 354
- constituent realism, 138
- constitutive micropsychism, 134
- constitutive panpsychism, 134
- content externalism, 336
- contingent physicalism, 193–212
  - Bailey's argument, 201–12
  - ontological categories, 201, *202*, 208–9
  - organisms, depiction/understanding of, 203, 204, 206–7
  - problems with, 203–12
  - qualification of, 206n26
  - thought experiment, 202, 203, 204, 205–8

  via negativa, use of, 210, 211–12
  Merricks's argument, 193–201
  objections to, 194–201
Copenhaver, Rebecca, 313
Coradint, Antonella, 176
Corcoran, Kevin, 168, 298, 330
correlation
  mental/physical correlations, 294–8
  contingency of, 297–8
  neural correlates of consciousness, 289
  private correlations, 136
cosmic authority problem, 240
Cosmides, Leda, 236
Cucu, Alin C., 280
CWC *see* complex whole causation

D
de dicto beliefs, 62–3, 67–72, 231
de dicto knowledge, 104, 224, 225, *227*
de dicto modality, 214
de dicto reference, 197
demonstrative knowledge, causal chain analysis, 214–15
Dennett, Daniel, 4, 93, 95, 249–50
de re beliefs, 62–3, 67–72, 231
de re knowledge, 104, 224–5, *227*
de re modality, *229*
de re reference, 195–6, 197
de re rigid designators, 192
de re seemings, 221
Descartes, René, 295, 312–13, 324–5
Descartes' Zombie, 85
Des Chene, Dennis, 183
descriptivism, 195
de se knowledge, 226
determinate concept possession, 219
developmental SD research projects, 322–32
  embodiment, logical space of, 322–8
  nature of the physical, 331–2
  psychological development, 330–1
diachronic determinism, 245
diachronic identity, 144, 145, 178, 181
  simple and complex views of, 155–6, 199–200
  see also personal identity
dichotomistic vocabulary, 255
directedness, 97–8
directly referential terms, 195–6
direct reflective self-awareness argument, 103–5
  fulfilled SD seemings, 104
  Husserlian fulfilment structures and self-knowledge, 103
  knowledge and intentionality, 103–4
direct self-awareness account (DSA), 73, 92–103
  causal-acquaintance principle, 100
  directedness, 97–8
  dualist seemings, 92–3
  explanatory conditions of, 101–2
  explanatory unity, 102–3
  giveness, 97–8
  implication of, 116
  knowledge by acquaintance, 94–5
  monadic intentionality, 96–7, 101
  self-awareness necessity, 99–100
  self-presenting property of direct self-awareness, 98–9
disembodiment, intuitions of, 194
dispositional theory, 342
dispositions, 83, 342
distribution
  phenomenal unity of consciousness, 123
  problem of, 128
diversity, dualist seemings, 90, 91, 101
divine, knowledge by acquaintance, 332
DNA, 184, 185–6
downward causation (DC), and physicalism, 177–8
DPA *see* dual process account
DSA *see* direct self-awareness account
dualism
  Aristotelian-style dualism, 182
  dichotomistic vocabulary, 255
  disembodiment intuitions, 194
  selfhood, 257
  and soul's existence, 284
dualist seemings, 84–93
  absence of, 107–8
  artificial intelligence test seemings, 90
  cognitive studies, 86–8
  direct self-awareness account, 92–3
  explanatory conditions of, 90–2, 101–2
  explanatory gap seemings, 84–5
  explanatory unity, 102–3
  first-person dualist seemings, 88–90
  and Moorean facts, 112–13
  neuroscientific explanations for, 110
  unity and simplicity seemings, 85–6
dual process account (DPA), 110–16
  dualist rival to DPA, 113–14
  dualist seemings and Moorean facts, 112–13
  generality problem for, 114–15
  scanning-monitoring approach, 115–16

E
easy meta-problem of consciousness, 83
Edwards, Jonathan, *Dictum*, supervenience argument, 263, 264, 268
EEG (electroencephalogram), 347–8, 349
ego, 286
eidetic intuition, 225–6, 230
electroencephalogram (EEG), 347–8, 349

electron analogy, subject necessity, 54–5
elimination, Barnett's simplicity argument, objection to, 76–7
embodiment, 322–8
  bodily soul view, 325–6
  conditions for, 328–9
  embodied and not-embodied, *300*
  embodied privacy, 89
  embodiment distinctions, 322–4
  importance of for SD, 328–30
  staunch hylomorphic SD, 326–8
  and unity of the person, 324–5
emergence
  consciousness, 39–40, 115, 252, 255, 256
  Hasker, 154
  soul, 283
  top-down causation, 264–70
emergent properties (EPs), 26, 29–41, 131–2
  as brute facts, 37–8
  characteristics of, 32
  classification of, 34–5
  combination problem, 38–9
  consciousness, emergence of, 39–40
  criticisms of, 35–41
  emergence and empirically equivalent theories, 36
  "emergence" as just a label, 35–6
  emergent vs. structural supervenience, 30–1
  ex nihilo problem, 36
  and faint-hearted naturalism, 32–5
  metaphysical aspects of, 31–2
  notion of, 31
  novel properties and new substances, 40–1
  sorites problems, 38
  theism, Bayesian argument for, 39–40
  twin hallmarks of emergence, 31–2
  weak and strong emergence, 31
  weak vs. strong property supervenience, 30
emergent substance dualism, 315
emergent supervenience (ES), 30–1, 268, 269n54, 270
empathy, and mirror neurons, 288
empirical equivalence, strong and weak, 289
empty intention, 104, 105, 106
endurance, mereological argument, 147
energy, conservation law, 353–4
ensouled animalism, 129
entailment, 98
entropic brain hypothesis, 350–1
entropy, 350, 351
EPA *see* extended physical atom view
epiphenomenalism, 236, 248, 268, 270, 295
epistemic fundament, 211, 212
epistemology
  applied SD research projects, 334–5
  basicality, 20
  combinatorial modes of explanation, 25–6
  Grand Story, 27–8
  human intellect, origin and nature of, 24
  naturalness, 20–1
  Ockham's Razor, 21–3
  ontological parsimony, 22–3
  questions, 290
  scientism, 25
  third-person ways of knowing, 25
EPs *see* emergent properties
ES *see* emergent supervenience
essentially grounded, 151, 206
Evans, C. Stephen, 294, 325
event-causation, 242
event–event causation, 241, 293
evolution, 27, 33, 236, 240, 338
Ewing, A. C., 161–2
exclusion argument *see* supervenience argument
ex nihilo problem, 36
experience, relational view of, 196
experiential conceptualism, Hopp, 196–7
explanatory gap seemings, 84–5
  cognitive science studies, 110–11
  intuitions, 112, 113–14
extended physical atom (EPA) view, 169–72
  rejection of, 171–2
externalism
  content externalism, 336
  and physicalism, 334–5
external relations, 47, 150, 227
extreme rational essentialism, 340

F
faculties
  attribution of to brain not soul, 287
  and capacities, 287–8
  "faculty" term/concept, 222, 287
faint-hearted libertarian freedom, and free will, 257–61
faint-hearted naturalism, 17–18, 27, 29, 32–5, 249
Fales, Evan, 7
first law of thermodynamics, 353, 354
first-person approach/perspective, 113
  consciousness, 50–3, 293
  dualist seemings, 88–90
  neuroscientific methodology, 292–3
  SD research projects, 319–20
first philosophy, argument against, 25, 291, 292
fission problem, personal identity, 166–7
fMRI data/studies, 52, 64, 294, 347, 349, 350
Fodor, Jerry, 92
folk epistemology, 216
free agents, as substances, 241–3
free will *see* libertarian free will
fulfilment structures, 157–8, 162–3
  Husserl, 103, 224–5

Fumerton, Richard, 55, 67
functional magnetic resonance imaging (fMRI), 52, 64, 294, 347, 349, 350
functional parts, 128
functional reduction principle, 129
functional roles, 200
  car engine analogy, 207
  mereological aggregates, 204, 205
functional unity, 48, 129
function argument, 128–31
function problem, 128

G
gambling study, regret correlation with orbitofrontal cortex, 69
Ganssle, Greg, 281, 282
generalization argument, mereological hierarchy, 155
general relativity theory, and conservation laws, 278
genes, 185
genuine wholes, 46–50, *49*, 56, 128
Ghirardi–Rimini–Weber (GWR) model, 355
Ginet, Carl, 238
giveness, 97–8
global naturalists, 17n6
God, 261, 300, 301, 305, 306–7
  soul relationship to, 282
  as wholly immaterial, 200–1
Goetz, Stewart
  awareness of spiritual self, 156, 157
  modal argument for SD, objection to, 230
  the physical, 331
Goff, P., 139
gold, substance characterization, 304–5
goodness, intuition of, 109
Goodwin, Brian, 186
gradualism problem, personal identity, 166
Grahek, Nikola, 71–2
Grand Story, précis and elements of, 27–8
GWR-spontaneous collapse theory, 355

H
Haldane, John, 64
hard meta-problem of consciousness, 82–92
  commonsense substance dualism, cognitive science studies, 86–8
  dualist seemings, 84–92
  unity and simplicity seemings, 85–6
Hart, W. D., 13, 353
Hasker, William, 132–3, 154, 160
healthcare, SD research projects, 337–9
Heisenberg's uncertainty principle, 348
high-road cognitive processes, 111, 112
history of science, analogies from, 70–1
Hoffman, Donald, 52
Hoffman, Joshua, 209–10
Hogan, Krista and Tatiana, 140
holenmerism, 282
holistic organicism, 289
  see also Organicism
holistic unity, 48, 128
Holstrom, Nancy, 244
homeodynamic systems, human persons as, 172–87
homogeneous continua, 152
Hopp, Walter, 92, 104–5
  consciousness, 286
  de re reference, 195
  experiential conceptualism, 196–7
Horst, Steve, 70
HUC *see* human unobservability concerns
Hudson, Hud, 207
Hull, David, 338
'human nature', 338
human unobservability concerns (HUC), 344–5
Hume, David
  phenomenological report, 105–6
  the self, unobservability of, 345n1
Husserl, Edmund
  de re knowledge, 104
  epistemology of defeasible modal knowledge, *227*
  fulfilment structures and self-knowledge, 103, 157–8
  intentionality, 195–6
  intuitive awareness, 230–1
  modal knowledge theory, 223–9
  ontology, 43–50, 226–7
  aggregates, 46–50, *49*, 56
  genuine wholes, 46–50, *49*, 56
  "laws of essence", 49
  part-whole distinctions, 44
  relations, 46–50
  separable and inseparable parts, 44–6
  pure perceptual states, 197n17
Hutteman, Andreas, 152
hylomorphism, 8–9, 79
  staunch hylomorphism, 179–87, 203, 204, 315, 326–8, 338
  thin particular hylormorphism, 181, 186–7

I
IA *see* impossibility argument
ideal gas law, "covering law" model of explanation, 293–4
identity *see* personal identity
IM *see* individual maximality
immanent causation (IC), 167–9, 243, 244
immaterial objects, *via negativa* characterization of, 210
immaterial self, 143

immaterial substance, as mind, 295
immediately based knowledge, 104
impossibility argument (IA), Bailey's
    objection to, 305–7
inclusion, 98
independence, and SD, 312, 313, 314
indeterminate causation, 260
indexicals, content and character of, 223
individual maximality (IM), 77, 207
indivisibility, personhood, 90
inertial motion, 20
infants, attribution of intentions to persons, 86
infinite regress problem, 244
information
  definition problem, 184
  as *sui generis*, 289
inseparable parts, 44–6, 228
integrative dualism, 328
integrative substance dualism, 315
integrity, Barnett's simplicity argument,
    objection to, 79–81
intellect, origin and nature of, 24
intensional fallacy objection, to argument from
    introspection, 60–7
  replies to
  absurd introspection skepticism, 63–6
  equivocation of propositional knowledge and
    knowledge by acquaintance, 62–3
  metaphysical nature of argument from
    introspection, 66–7
intentionality, 88
  and *de re* reference, 195–6
  and free agency, 238
  Husserl, 195–6
  knowledge and, 103–4, 224
  as monadic property, 96–7, 101
  of sensations, 217
interactive dualism, 353
internalism
  epistemology, 334–5
  and physicalism, 93n78
internal relations, 137–8, 227, 229
  Husserlian ontology, 46–7
  SD research projects, 318
  systems theory, 175
introspection, argument from, 59–73
  intensional fallacy objection, 60–7
  replies to, 62–7
  neurological evidence objection, 67–72
  Searle's argument against, 72–3
introspection, and free agency, 239, 240
introspective awareness, 239, 240, 260
introspective skepticism, 68–9
intuitions
  and *a priori* knowledge, 221
  and beliefs, 109–10
  categorical intuition, 225
  disembodiment, 194
  dispositions, 83
  eidetic intuition, 225–6, 230
  evidential nature of, Bealer, 219, 220, 221–2
  explanatory gap, 112, 113–14
  first-order properties, 109
  modal intuition, 197–8, 232
  rational intuition, 212, 213–29
  substance dualism, 8–9
intuitive awareness, 230–1
intuitive knowledge, 218
intuitive seemings, 221
"I" term
  as *de re* rigid designator, 192
  lack of conceptual/semantic meaning, 222–3

J
Jackson, Frank, 28, 174
Jacobs, Jonathan D., 235, 236, 261–2
James, William, 3, 134
Jedlicka, Peter, 357–9
Junky Worlds, 151–2

K
Kahneman, Daniel, 111
Kane, Robert, 257–61
Kaplan, David, 223
Kelly, Thomas, 91
Kim, Jaegwon, 22, 29, 33, 35
  argument from introspection, objection to,
    59–60
  causal pairing argument, 280, 281–2
  closure principle, 248
  consciousness and material/immaterial
    substance, 302
  generalization argument, 155
  mental causation model, 356–7
  neuroscientific correlations, 297
  personhood, 74
  physicalism, 18–19, 271
  qualia supervenience, 64, 69, 110, 297
  scanning-monitoring approach, 115–16
  self-awareness failure, 109
  strict physicalism, 148
  supervenience argument, 263–4
knowledge
  a priori knowledge, 213, 221
  by acquaintance, 62–3, 94–5, 214, 332
  causal chain analysis, 214–15, 217, 218
  causal modal knowledge, 215
  de dicto knowledge, 104, 224, 225, *227*
  de re knowledge, 104, 224–5, *227*
  de se knowledge, 226
  as fulfilment, 224
  and intentionality, 103–4, 224

  intuitive knowledge, 218
  and modal facts, 215
  ostensive definitions, 51
  propositional knowledge, 62–3
  scientific knowledge, 25
  self-knowledge, 88–90, 95, 103, 226
  see also modal knowledge
Koons, Robert, 39–40, 179, 180
Kriegel, Uriah, 135–6
Kripke, Saul, 219, 220, 233

L
Langsam, Harold, 11, 320, 321
Lavazza, Andrea, 19
law, philosophy of, 335
"laws of essence", Husserlian ontology, 49
laws of nature
  mereological hierarchy, 153–5
  violation of, 346
Leibniz, Gottfried Wilhelm, 119–20, 328
Leibniz's Law, 136
Leibniz's Mill, 129, 305
Levitin, Daniel, 81
Lewis, C. I., 125
Lewontin, Richard, 186
libertarian free will (LFW), 235, 289
  causality, 259–62
  ontology of, 260
  commonsense understanding of, 240
  compatibilist theory, 239–40
  consciousness, 252–4
  deciding and acting, phenomenological difference between, 238–9
  Dennett, 249–50
  incompatibilist approaches, 240–1
  intentionality, 238
  justification for, 237–40
  minimalist naturalism and, 261–2
  naturalistic constraints, 250–1, 257–9
  necessary conditions for, 258
  neurobiological reality of, 253
  Searle, 246–7, 250–7
  self-awareness, 238
life, origin of, 27
life principle, 130
likelihood principle, 296
linguistic dependence, self-awareness arguments, objection to, 107
linguistic reduction, 94
local naturalists, 17n6
location, experience of, 196
location problem, naturalist ontology, 28
logical space of substance dualism, 312–16
  historical argument, 312–13
  substance argument, 314
Lowe, E. J., 43, 53, 351–3, *352*
low-road cognitive processes, 111, 112
Lund, David, 162–3
Lycan, William, 6, 275–6, 286
Lyon, William, 4

M
McGinn, Colin, 25–6, 36, 336
McQueen, K. J., 354–6
macro-identity, 200
macro-objects, 266–7
Madden, Rory, 79
Madell, Geoffrey, 142, 217
magic, 305–7
MAR *see* Metaphysical Aristotelianism
Markosian, Ned, 132
Marmodoro, Anna, 172, 326
Martin, Michael, 7
MAs *see* mereological aggregates
materialism, 3, 6, 106
material objects, *via negativa* characterization of, 210
matter, nature/properties of, 29–30, 36
Maudlin, Tim, 353–4
MaxCon view, 132, 146
maximality, Barnett's simplicity argument, objection to, 77–8
Mayr, E., 338
mechanical philosophy, 318
mechanistic question, mental and physical entity causal interaction, 276–7, 306, 307
Medieval Aristotelians, 164, 168, 181, 183
Meixner, Uwe, 41, 101
memories, induction by brain stimulation, 140
mental causation, *247*
  Lowe's model of, 351–3, *352*
mental health, SD research projects, 337–9
mental holism, 93n77
mentality, 96
mental/physical correlations, 294–8
  contingency of, 297–8
mental powers essentialism, 89
mental privacy, 89
mental properties
  causality, 268–9
  characteristics of, 169–70
  instantiation of, 54, 55
mental-sharing
  denial of, 139
  empirical cases, 140
  and subjectivity, 135–6
mental states
  attribution of intentions, 86
  and brain states, 67–8, 247–8, 359
  causal reduction, 256
  conjoined twins, 140
  and consciousness, 55

  features of, 113–14
  intrinsic nature of, 64
  and introspection, 63–4
  neuroscientific correlations, 297
  physical nature of, 65
  Ravenscroft, 301–2
  reasons-giving relations, 196–7
  self-presenting property, 99
mereological aggregates (MAs), 145, 147–8
  contingent physicalism, Bailey, 204–5
  functional roles, 204, 205
  as mereological inconstant objects, 163–87
  circularity problem, 165–6
  defeaters for complex view of personal identity, 165–9
  fission problem, 166–7
  gradualism problem, 166
  immanent causation, 167–9
  ontology of MAs, 164–5
  sorites problem, 166
  transitivity problem, 166
mereological complexity, *50*, 108
mereological essentialism, 144–87
  human persons as enduring, simple souls, 155–63
  awareness of our spiritual self, 156–8
  diachronic personal identity, 155–6
  fulilment structures, 162–3
  personal identity, 155–6
  Platinga's replacement argument, 158–9
  rational deliberation, 160–3
  teleologically arranged action plans, 162–3
  teleological reasoning, 160–2
  mereological aggregates as inconstant objects, 163–87
  mereological argument, 145–6
  clarification of, 147–8
  mereological hierarchy, 148–55
  physicalist alternatives to human persons as MAs, 169–87
  atomic simples, 169–72
  homeodynamic systems, 172–87
  simple souls, human persons as, 155–63
  simplicity, 146–7
  systems approach (and objections to), 172–87
  complex systems as enduring continuants, 178–9
  complex systems as synchronically unified wholes, 175–7
  downward causation, harmonization with physicalism, 177–8
  spooky entities, 174–5
  staunch hylomorphism superiority over systems theory, 179–87
  what we seem to know about ourselves, 174
mereological hierarchy (MH), 148–55
  category of individual, 149
  category of property/relation, 149–50
  causation, 153
  concepts and laws, 153–5
  general depiction of, 149
  generalization argument, 155
  naturalist ontology, 29
  structure of, 150–2
mereological simplicity, 45–6, *50*, 132, 146–7
  Barnett's simplicity argument, 75–82
  Husserl, 81
  souls, 186–7
mereological supervenience, 150
mereological universalism, 47n19
mere substance dualism thesis, 13–15
Merricks, Trenton, contingent physicalism and the modal argument, 193–201
Metaphysical Aristotelianism (MAR), 182–4
methodological questions, 290
methodological SD research projects, 316–22
  consciousness-first approach, 320–2
  empirically informed and testable SD, 316
  first-person approach, 319–20
  parts, wholes, and unity, 318
  person/subject, 319
  philosophy of mind, 316–19
  properties, powers, and substances, 317
  teleology, 318–19
MH *see* mereological hierarchy
micro-causation, 152
micro-determinism, 152
micro-government, 152
microphysicalism, 152
microphysical priority, naturalist ontology, 29
mind
  and brain, 295
  as immaterial substance, 295
mind-body problem, 6, 7
mind emulation, 341
mindfulness, 333
mine-ness
  consciousness states, 88
  necessity of, 137, 143
mirror neurons, 288
modal argument for SD, 191–233
  clarification of argument, 192
  conceivability and possibility, 212–13
  defense of the argument, 193–229
  contingent physicalism, 193–212
  objections to, 230–3
  circularity of modal argument, 230–1
  dualist thought experiments, problems with, 232–3

  duality of concepts issue, 231
  epistemic and metaphysical possibility, confusion between, 233
  physicalist and dualist intuitions, cancelling out of, 231–2
  statement of argument, 191–2
  see also modal knowledge
modal beliefs, 213, 215
modal errors, 216, 219, 220
modal facts, and knowledge, 215
modal fallibility, 218
modal intuitions, 197–8, 232
modal knowledge, 213–29
  Bealer's argument, 218–23
  critique of, 221–3
  intuitions as evidence, 219, 220, 221–2
  modal errors, 220
  Husserl's theory, 223–9
  dependence and necessity, ontology of, 229
  epistemology of defeasible modal knowledge, *227*
  fulfilment structures, 224–5
  knowledge and intentionality, 224
  ontology of *de re* modality, *229*
  self-knowledge, 226
  O'Connor's argument, 213–18
  causal modal knowledge, 215
  de dicto modality, 214
  modal error, sources of, 216
  rational intuition, 214, 216
modes, phenomenal states, 138
molecule-in-a-wheel case, Steward, 151
moments, Husserlian ontology, 44
monadic intentionality, 96–7, 101
monadic property-instance, 45
monism, 27
monotheistic physicalism (MP), 300–1
monotheistic substance dualism (MSD), 300–1
Monton, Bradley, 299
Moorean facts, 92, 112–13
Moore, G. E., 109
MP *see* monotheistic physicalism
MSD *see* monotheistic substance dualism
Mumford, Stephen, 239
Murphy, Nancey, 153, 154, 172–6, 285–6
mysterian naturalism, 36
mysteriousness, soul and, 360
"my" term, as *de re* rigid designator, 192

N
Nagel, Thomas, 240
natural function, and teleology, 246
naturalism, 7–8, 9, 16–42, 148, 234
  causal relations, 242–3
  emergence of, 29–41
  epistemic issues, 20–3
  basicality, 20
  naturalness, 20–1
  Ockham's Razor, 21–3
  ontological parsimony, 22–3
  faint-hearted naturalism, 17–18, 27, 29, 32–5, 249
  faith in, 6–7
  free will, 251–4, 261–2
  mysterian naturalism, 36
  ontology, 28–9
  location problem, 28
  microphysical priority and mereological hierarchy, 29
  scientific naturalism, 16–20, 23–9
  and standard picture of philosophy, 18–20
  and staunch libertarian agency, 235–7
  staunch naturalism, 17, 235, 249, 262
  worldview of, 19
naturalist free will, *252*
naturalistic causation, 244
naturalistic substance dualism, 333–4
naturalist mereological hierarchy, *149*
naturalness, of entities, 20–1
natural theology, consciousness-first approach, 333
natural world, layered view of, 148–55
  see also organisms
NCC *see* neural correlates of consciousness
near-death experiences (NDEs), 90, 333
necessitation, scientific explanation, 294
nervous system, 357
neural contingency, 89
neural correlates of consciousness (NCC), 289
neural networks, 357–8
neurobiological determinism, 253
neurological evidence objection, to argument from introspection (and replies to), 67–72
  Grahek's challenge to phenomenal essence of pain, 71–2
  history of science, dubious analogies from, 70–1
  neuroscientific loss of introspective skepticism, 68–9
  physicalism, illicit presumption of, 70
neurological theses, 323–4
neuronal signals, 348, 349, 358
neuroscience, 108, 249
  consciousness, defining, 51, 52
  first-person reports, 292–3
  introspective skepticism, loss of, 68–9
  and philosophy of mind, 290–2
  and the soul, 347
  see also neuroscientific objections to SD
neuroscientific correlations, 294–8
  contingency of, 297–8
neuroscientific objections to SD, 284–301

- empirically equivalent findings/theories of neuroscience, 288–9
- explanatory impotence of SD, 285–8
- first-person reports, reliance on, 292–3
- mental/physical correlations, 294–8
- necessitation, 294
- parsimony objection, 298–301
- philosophy of mind, issues in, 290–2
- scientific explanation, 293–4
- soul-of-the-gaps argument, 284–5

Nijhout, H. F., 186
Noonan, Harold, 167
Norden, Jeanette, 347

O

objections to substance dualism, 3–5, 6, 275–307
- Bailey's magical mystery tour objection, 305–7
- causal interaction problems, 276–83
- neuroscientific objections, 284–301
- Ravenscroft's physicalist defeater, 303–5
- soul, explanatory impotence of, 301–7

objects
- compatibility/incompatibility, 228–9
- stress-energy tensor, 278

objectual phenomenal unity, 143
obsessive-compulsive disorder (OCD), 331
Ockham's Razor, 21–3
O'Connor, Timothy, 36, 41, 100, 234
- modal knowledge, 213–18
- non-reductive physicalism, argument against, 267–70

Oderberg, David, 168
Okon, Elias, 353–4
Olson, Eric, 129–30, 302
ontological emergence, 115
ontological parsimony, 22–3, 33
ontology
- Bailey's categories, 201, *202*, 208–9
- Husserlian, 43–50, 226–7
- aggregates, 46–50, *49*, 56
- genuine wholes, 46–50, *49*, 56
- "laws of essence", 49
- part-whole distinctions, 44
- relations, 46–50
- separable and inseparable parts, 44–6
- philosophy of mind, 316–19
- questions, 290
- standard naturalist ontology, 28–9

orbitofrontal cortex, regret correlation, 69
ordered aggregates, *49*
Organicism, 182, 183–4, 185, 186
- *see also* holistic organicism

organisms
- Bailey's depiction/understanding of, 203, 204, 206–7
- theories of, and staunch hylomorphism, 179–87

organocentric view, 185–6
overlapping parts view, 130–1, *131*
Owen, Matthew, 289

P

Page, Ben, 39
pain
- affective-motivational aspect of, 72
- nature of, 71
- perception of, 221
- phenomenal essence of, 71–2

Palette Problem, 38–9
panpsychism, 10–11, 12, 20
- contra phenomenal unity of consciousness, 133–41

panpsychist phenomenal unity, 134
- privacy argument against, 135–9

PAP *see* principle of alternative possibilities
paper airplane analogy, emergent properties, 154
Papineau, David, 24, 29, 33–4, 84, 154, 291
parsimony objection, neuroscientific, 298–301
parthood, 200
- Bailey's criterion for, 204–5
- Husserlian ontology, 44–6
- SD research projects, 318

part-whole distinctions, Husserlian ontology, 44
Pasnau, Robert, 182
passive liabilities, 243
passive thoughts, 100
Penfield, Wilder, 140, 240
perceptions
- Hume, 345n1
- pure perceptual states, Husserl, 197n17

perceptual demonstratives, 196
perdurance, 147
persistence, varieties of, 147
personal identity (PI)
- defeaters for complex view of, 165–9
- circularity problem, 165–6
- fission problem, 166–7
- gradualism problem, 166
- immanent causation, 167–9
- sorites problem, 166
- transitivity problem, 166
- grounding of, 199
- simple and complex views of, 155–6, 199–200
- sortal dependency of PI, 167
- and soul identity, 199

person emulation, 341
personhood, 53–4, 74
- indivisibility, 90
- SD research projects, 319

  unity of person, 324–5
person-pair fusion, 76, 77
person, unity of, 324–5
PET scans, 347, 349
phenomenal holism, 122
phenomenal qualia, 69, 109
phenomenal states
  content of, 127
  internal relations, 137–8
  mine-ness necessity, 137
  modes, 138
  neo-Aristotelian ontology of, 137–9
  self-presentation, 136–7, 138–9
  subjectivity, 135–6
  subject necessity, 54–5
phenomenal transparency, 11–12
phenomenal unity of consciousness, 89, 119–43
  contra physicalism, 123–33
  anti-distribution argument, 123–4
  emergentist objection, 131–2
  emergent substantial souls, 132–3
  function argument, 128–31
  micro-subject problem, 127–8
  no phenomenal processing problem, 127
  process argument, 126
  temporal distribution problem, 125–6
  contra Russellian panpsychism, 133–41
  combination problem, 134–5
  privacy argument against panpsychist phenomenal unity, 135–9
  general objections to, 141–3
  Bayne's explanatory failure objection, 142–3
  rejection of phenomenal unity, 141–2
  process problem, 124, 126
  rejection of, 141–2
  subject phenomenal unity, 122
  temporal distribution problem, 125–6
  what-it-is-like, 121
phenomenological report, Hume's, 105–6
philosophy of law, 335
philosophy of mind
  consciousness-first approach, 11, 320–2
  neuroscience and, 290–2
  rhetorical devices used in, 5
  SD research projects, 316–19
  thought experiments, 85–6
  and worldviews, 16–17
philosophy of religion, 332–4
philosophy of science, 70
photons, nature of, 209
physical
  nature of, 331–2
  opacity of, 12–13
physicalism, 29, 66, 116n166
  arguments against, 12
  arguments for, 110
  atomic simples, human persons as, 169–72
  contra phenomenal unity of consciousness, 123–33
  dichotomistic vocabulary, 255
  distribution thesis, 123
  and downward causation, 177–8
  externalism, 334–5
  falseness of, 74, 84
  homeodynamic systems, human persons as, 172–87
  illicit presumption of, 70
  and internalism, 93n78
  mental states, 65
  monotheistic physicalism, 300–1
  opacity of the physical, 12–13
  process argument, 126
  process problem, 127
  Ravenscroft's objection to SD, 303–5
  standard/strong physicalism, 9–13, 18
  and staunch libertarian agency, 235–7
  strict physicalism, 148
  strong physicalism, 17, 27
  weak physicalism, 17–18, 27, 29
  see also contingent physicalism
physical objects (POs)
  Bailey's ontology, 201, *202*
  modal argument for SD, 191–2
  as not wholly physical, 200
  proper constituents, 198
  see also complex objects
physical persistence, 89
PI *see* personal identity
pieces, Husserlian ontology, 44
Pinker, Steven, 52
Pitts, J. Brian, 280
Place, U.T., 3
Plantinga, Alvin, 158–9
Plato, 312, 313
PM *see* property maximality
Pointy View, 132, 133, 146
Popper, Karl, 6
POs *see* physical objects
positron emission tomography (PET), 347, 349
possibility, modal argument for SD, 212–13
possible-worlds, 198
powers substance, 314
primary consciousness, 350
primitive consciousness, 96
principle of alternative possibilities (PAP), 258
priority microphysicalism, 152
privacy argument, against panpsychist phenomenal unity, 135–41
  mental-sharing and subjectivity, 135–6
  mine-ness necessity, 137

- objections to privacy argument, 139–41
  - denial of mental-sharing, 139
  - rejection of privacy, 139–41
  - phenomenal states, neo-Aristotelian ontology of, 137–9
  - self-presentation, 136–7
- private correlations, 136
- process problem
  - phenomenal unity of consciousness, 124, 126
  - of physicalism, 127
- properties
  - causal theory of, 268–9
  - SD research projects, 317
- property dualism, 10, 110, 230
- property-exemplification event identity, 200
- property maximality (PM), 78, 207
- property reduction principle, 126, 132
- propositional knowledge, 62–3
- psychedelic drugs, consciousness states, 350, 351
- psychic energy, 353
- psychological development, SD research projects, 330–1
- psychological persistence, 89
- psychological precondition, 211–12
- psychophysical correlations, 294–8
- pure perceptual states, Husserl, 197n17
- Putnam, H., 71, 217
- Pyysïäïnen, Ilkka, 107

Q

- qualia
  - introspection, 69
  - phenomenal qualia, 69, 109
- qualia supervenience, 64, 69, 110, 297
- qualitative parsimony, 298, 299
- quantitative parsimony, 298
- quantum entanglement, 279
- quantum state, 355n25
- quantum stimuli, brain responsiveness to, 357–9
- quantum theory, top/down causation, 257

R

- Rachels, James, 236
- Rasmussen, Joshua, 296
- rational intuitions, 212, 213–29
- Ravenscroft, Ian, explanatory impotence of SD, 285, 301–2, 311
- reasons, as teleological ends, 261
- reasons explanations, 256–7
- regret, gambling study, 69
- Reid, Thomas, 53–4
- Reinach, Adolf, 211
- relations, Husserlian ontology, 46–50
  - *see also* external relations; internal relations
- relativity theory, and conservation law, 278
- religion
  - philosophy of religion and SD, 332–4
  - and substance dualism, 91, 108
- replacement argument, Plantinga's, 158–9
- research programs for substance dualism, 311–43
  - applied research projects, 332–42
  - conceptual research projects, 312–16
  - developmental research projects, 322–32
  - methodological research projects, 316–22
- rigid essentially grounded, 151
- rigid existentially grounded, 206
- Robinson, Howard, 19, 324
- robust concept possession, 219
- Rodrigues, José Gusmão, 6, 275
- Rosenberg, Gregg, 52
- Rosenkrantz, Gary S., 209–10
- Russell, Bertrand
  - knowledge by acquaintance, 94
  - knowledge, ostensive definitions, 51
  - panpsychism, contra phenomenal unity of consciousness, 133–41
- Ryle, Gilbert, 4

S

- salted water analogy, soul, 327
- scanning-monitoring approach (SMA), 115–16
- scattered parts view, 123–4, *124*, 127
- scattered wholes, 125
- Schneider, Susan, 341
- Schrödinger equation, 355n25
- Schwartz, Jeffrey M., 331, 349–50
- Schwitzgebel, Eric, 51
- sciences
  - authority of, 292
  - reductions in, 70
  - and soul's redundancy, 285–6
- scientific explanation, 293–4
- scientific knowledge/theories, 25
- scientific laws *see* laws of nature
- scientific naturalism, 16–20, 23–9
  - epistemic attitude, 24–6
  - combinatorial modes of explanation, 25–6
  - Grand Story, 27–8
  - human intellect, origin and nature of, 24
  - scientism, 25
  - third-person ways of knowing, 25
- scientism, 25
- Scruton, Roger, 284
- SE *see* strong emergence
- Seager, William, 284
- Searle, John, 6–7, 67, 141
  - argument from introspection, objection to, 72–3
  - Chinese Room thought experiment, 86

  free will, 246–7, 250–7
  materialism, 106
  philosophy of mind, rhetorical devices, 5
  physics, SD's incompatibility with, 278
secondary consciousness, 350
the self, 8, 106
  first-order awareness of, 109
  Hume, 345n1
  realism about, 53–4
  simple self, 75, 89
self-awareness, 74–116
  awareness of spiritual self, 156–8
  direct self-awareness account, 73, 92–103
  free will, 238
  genuine self-awareness, 115, 116
  knowledge by acquaintance, 95
  mereological essentialism, 156–8
  objections to self-awareness arguments, 105–16
  absent dualist seemings, 107–8
  commonsense materialism, 106
  dual-process account, 110–16
  Hume's phenomenological report, 105–6
  linguistic and theoretical dependence, 107
  neuroscience studies, 108
  religious etiology, 108
  Searle's argument, 72–3
  self-awareness failure, 108–10
  self-awareness necessity, 99–100
  simplicity argument, Barnett's, 75–82
  temporal self-awareness, *156*
self-forming action (SFA), 258
selfhood, dualism, 257
self-intuitions, 83
self-knowledge, 88–90, 95, 103, 226
self-presentation, 136–7
self-presenting property, 73, 98–100
Sellars, Wilfrid, 25
semantic questions, 290
semantic reduction, 94
sensations, intentionality of, 217
separable parts, 44–6, 228
serial bonding, 151
series-of-instantaneous-entities view, 147
SFA *see* self-forming action
Sharpe, Kevin, 298
Sider, Theodore, 265, 266
Siegel, Daniel, 249, 291
simple dualism, 322, 323
simple self, 75, 89
simple souls (SSs), 132, 169–71, 286
  mental subjects as, 170–1
  mereological essentialism, 155–63
  modal argument for SD, 191–2
simplicity
  Husserlian ontology, 44–6
  mereological simplicity, 146–7
simplicity argument, Barnett's, 75–82
  objections to, 76–81
  elimination objection, 76–7
  integrity objection, 79–81
  maximality objection, 77–8
  substance objection, 78–9
  statement of argument, 75–6
simplicity seemings, 85–6
simulacrum hypothesis, van Inwagen, 197
skepticism, absurd introspection skepticism, 63–6
Skrbina, David, 134
SLA *see* staunch libertarian agency
SMA *see* scanning-monitoring approach
Smart, J. J. C., 26
social contract, 250
social epistemology, 335
social psychology, 337
sorites problems
  emergent properties, 38
  personal identity, 166
sortal dependency, of personal identity, 167
soul, 40, 74, 116n165, 143, 181
  bodily soul view, 325–6
  brain interactions, 295, 344–60
  concept of, 209, 302, 347
  Corcoran's objection to, 330
  Descartes, 324–5
  emergence of, 283
  emergent substantial souls, 132–3
  ensouled animalism, 129
  essence of, 330–1
  features of in SD theory, 303
  folk psychology of souls, 107
  human unobservability concerns, 344–5
  location of, 282, 283
  mereological argument, 145–6
  as mereologically simple, 186–7
  Metaphysical Aristotelianism, 183–4
  and mysteriousness, 360
  and personal identity, 199
  salted water analogy, 327
  soul-of-the-gaps argument, 284–5
  soul stuff, problematic nature of, 302
  spatial location of, 282, 283
  structure of, 330
  tethered body/tethered soul, 306, 307
  Thomisistic dualists' claim about, 283
  unintelligibility of, 209–10
  unobservability of, 344–5, 360
  see also simple souls
SP *see* structural property
spatial resolution, 348
Sperry, Roger, 252
spirit *see* soul
split-brain phenomena, 133, 141
spooky entities, 153, 174–5

SPS *see* strong property supervenience
SSs *see* simple souls
standard mereological hierarchy *see* mereological hierarchy
standard physicalism *see* strong physicalism
state–state causation, 293
staunch hylomorphic substance dualism, 315
staunch hylomorphism, 203, 204, 315, 326–8, 338
  superiority over systems theory, 179–87
staunch libertarian agency (SLA), 234–71
  alternatives to, 249–62
  creative compatibilism, Searle, 250–7
  faint-hearted libertarian freedom, Kane, 257–61
  free will irrealism, Dennett, 249–50
  minimalist naturalism and libertarianism, Timpe and Jacobs, 261–2
  causal closure of the physical, 263–70
  features of, 241–9
  free agents act for teleological ends, 245–6
  free agents are substances, 241–3
  free agents exhibit top/down causation, 246–9
  free agents as first movers, 243–5
  free agents have active power, 243
  free agents have categorical ability, 245
  formal characterization of, 241
  libertarian free will, justification for, 237–40
  metaphysical worldview implications, 240–9
  and naturalism/physicalism, 235–7
  statement of argument, 237
  top-down causation, 263–70
staunch naturalism, 17, 235, 249, 262
Steinhart, Eric, 7
Steward, Helen, 151, 153, 265
Stoljar, Daniel, 18
Strawson, Galen
  emergence as brute facts, 37–8
  physicalism, 17
  the self, 53
  subject necessity, 55
  subject phenomenal unity, 122
strength, dualist seemings, 91, 102
stress-energy tensor, of objects, 278
strict physicalism, 148
strong emergence (SE), 31
strong libertarian agency *see* staunch libertarian agency
strong physicalism, 9–13, 17, 18, 27
strong property supervenience (SPS), 30, 35, 336
structural property (SP), 31
structured parts view, 126
Suarez, Francisco, 181
subject complexity, 56, 120, 142
subjecthood, 12, 319
subjectivity, and mental-sharing, 135–6
subject necessity, 54–5, 88, 127
subject phenomenal unity, 122, 143
subject simplicity, 56, 120, 142
subject-summing problem, 134–5
substance, 142
  Barnett's simplicity argument, objection to, 78–9
  characterization of, 304–5
  contemporary substance dualism, 314–16
  free agents as, 241–3
  logical space of SD, 314
  material/immaterial substance, 302
  notion of, 304–5
  SD research projects, 317
  theory of, 303
substance causation, 242, 265
substance dualism (SD)
  in 21st century, 3–15
  Aristotelianism, revival of, 8–9
  as commonsense belief, 74, 86–8
  consciousness, 8
  consciousness-first philosophy of mind, 11
  direct reflective self-awareness argument for, 103–5
  direct self-awareness account, 92–103
  epistemic justification for, 116
  explanatory impotence of, 301–2, 303
  faith of anti-dualism, 6–7
  introspection arguments for, 59–73
  intuitions, 8–9
  logical space of SD, 312–16
  mere substance dualism thesis, 13–15
  modal argument for, 191–233
  naturalistic substance dualism, 7–8
  objections to, 3–5, 6, 275–307
  and religion, 91, 108
  resurgence of, 5
  scientific defeat of, 116
  support for, 4, 15
  worldview issues, 9
substance pluralism, 314
substantial agent, Kane's denial of, 259
substantial subjects, realism about, 53–4
substratum substance, 314
Sudarsky, Daniel, 353–4
sui generis property, 40
supervenience, 336
supervenience argument
  Edwards's *Dictum*, 263, 264, 268
  Kim, 263–4
Swinburne, Richard, 65, 293
  categorical/dispositional theory, 342
  embodiment conditions, 329
synaptic cleft, 81
synchronically unified wholes, complex systems as, 175–7

synchronic determinism, 245
synchronous supervenience, 264
System 1 and System 2 cognitive processes, 111
systems approach/theory, 172–87
  objections to, 173–87
  complex systems as enduring continuants, 178–9
  complex systems as synchronically unified wholes, 175–7
  downward causation, harmonization with physicalism, 177–8
  spooky entities, 174–5
  staunch hylomorphism's superiority over systems theory, 179–87
  what we seem to know about ourselves, 174

T

Tabaczek, Mariusz
  causality, 177, 178
  emergent properties, 32, 317
  hylomorphism, 180
Taliaferro, Charles
  consciousness, 10
  embodiment conditions, 328–9
  mereological essentialism, 144
  the physical, 331
technology, SD approaches to, 339–42
teleological functional unity, 48
teleological goals, free agents, 245–6
teleological holistic unity, 48
teleological reasoning, 160–2
teleological view, of reasons, 261
teleology, SD research projects, 318–19
temperature analogy, argument from introspection, objection to, 68
temporal distribution problem, 125–6
temporal resolution, 348
temporal self-awareness, *156*
tethered body/tethered soul, 306, 307
theism, 9, 39–40
thermodynamics, first law of, 353, 354
thinking immaterial object (TIO), 210–11, 212
thinking material object (TMO), 211
thin particular hylormorphism, 181, 186–7
third-person perspective, 25, 113
Thomistic dualists, and the soul, 283
thought
  passive and active thoughts, 100
  System 1 and System 2 processes, 111
thought experiments
  contingent physicalism, Bailey, 202, 203, 204, 205–8
  dualist thought experiments, 232
  introspective awareness, 239
  Leibniz's Mill, 129, 305
  role of, 290
time, composition of, 37
Timpe, Kevin, 235, 236, 261–2
TIO *see* thinking immaterial object
TMO *see* thinking material object
Tooby, John, 236
Tooley, Michael, 295
top-down causation, 153, 154, 256
  emergence problem, 264–70
  free agents, 246–9
  libertarian free will, 263–70
  and quantum gaps, 257
topological integration, 79
topological integrity, 81
topological scattering, 79–80
transhumanism, 340, 341, 342
transitivity problem, personal identity, 166
tree, sensation of, 217

U

ubiquity, dualist seemings, 91, 101
ultimate responsibility (UR), 258
uncertainty principle, Heisenberg's, 348
Unger, Peter, 81, 86
unity of the person, embodiment and, 324–5
unity, SD research projects, 318
unity seemings, 85–6
UR *see* ultimate responsibility

V

van Inwagen, Peter
  animalism, defence of, 129–30, 157
  modal intuitions, 232
  replacement argument, response to Plantinga, 158–9
  simulacrum hypothesis, 197
via negativa, Bailey's use of, 210, 211–12
virtual reality (VR), 342

W

wave function, 355n25
weak emergence (WE), 31
weak physicalism, 17–18, 27, 29
weak property supervenience (WPS), 30
Wellman, H. M., 87
Wells, Jonathan, 184
what-it-is-like, phenomenal unity, 121
what we seem to know about ourselves, 174
wheel of molecules, consciousness analogy, Sperry, 252, 255, 256
wholes
  genuine wholes, 46–50, *49*, 56, 128
  SD research projects, 318
wholly immaterial objects, 201–2
wholly material objects, 201–2
wholly and not-wholly material, *300*

Wiggins, David, 167
Willard, Dallas, 103
window-breaking analogy, causation, 265–6
wishful beliefs/thinking, 350
worldviews
  issues, 9, 290
  naturalistic worldviews, 17–18, 19
  philosophy of mind, 16–17
WPS *see* weak property supervenience

Y
Yandell, Keith, 277
Yang, Eric, 78

Z
Zimmerman, Dean, 13
zombies, 85, 89
Zuboffian Brain Separation, 86